Remarks

In view of the above amendments and the following remarks, reconsideration of the outstanding office action is respectfully requested.

Applicants submit that the first and second amendments to the specification rectify inconsistencies between the corrected formal drawings submitted herewith and the description of the drawings in the present application, the third amendment corrects a typographical error in referencing a particular example, and the remaining amendments correct two paragraphs to delete improper hypertext language.

Enclosed herewith are corrected drawings that overcome the objections identified in form PTO-948 that accompanied the outstanding office action. The objection to the drawings should therefore be withdrawn.

The rejection of claims 1-5, 11-15, 50, 53, 64, and 65 under 35 U.S.C. § 112 (2nd para.) for indefiniteness is respectfully traversed in view of the above amendments, except as noted in the remarks below.

With regard to claims 5, 15, and 53, the U.S. Patent and Trademark Office ("PTO") has taken the position that the recitation of "protein's activity is salt tolerant" is indefinite. Applicants respectfully disagree, because one of ordinary skill in the art would understand what is meant by a protein activity that is "salt tolerant." In particular, claim 5 depends from claim 4, which is drawn to an isolated heme-binding protein that is "isolated from *Halobacterium salinarium*." Claim 15 and 53, drawn to a blood substitute and chimeric protein, respectively, also specify that the bacterial heme binding protein is "isolated from *Halobacterium salinarium*." *Halobacterium salinarium* is an extreme halophile, a group of archaebacteria found in environments characterized as nearly NaCl-saturated (*see* Bandyopadhyay et al., "Salt Dependent Stability and Unfolding of [Fe2-S2] Ferredoxin of *Halobacterium salinarium*: Spectroscopic Investigations," *Biophysical Journal* 79:501-510, at 501, 1st full para. (2000) ("Bandyopadhyay") (attached hereto as Exhibit 1)), and it is well-known to those of ordinary skill in the art that proteins produced by halophilic bacteria require multimolar salt concentrations for activity (*see e.g.,* Jaenicke et al., "The Stability of Proteins in Extreme Environments," *Current Opinion in Structural Biology* 8:738-748 (1998) at pg. 744, rt. col., 1st full para. (attached hereto as Exhibit 2)). For example, Britton et al., "Insights into the Molecular Basis of Salt Tolerance from the Study of Glutamate

Dehydrogenase from *Halobacterium salinarium,*" J. Biol. Chem. 273(15):9023-9030 at 9023, 1st full para. (1998) (attached hereto as Exhibit 3) teaches that halophilic bacteria are known to live in environments where salt concentrations can exceed 3M, accumulate inorganic ions with the cell at concentration equivalent to or greater than that of the environment, and have proteins which are adapted to function under these high salt conditions. In addition, Bandyopadhyay teaches that *H. salinarium* proteins need a high salt concentration to maintain structural and functional activity (pg. 504, rt. col., 1st full para.), and that the secondary structure of a *H. salinarium* protein was found to be stable in salt concentrations from greater than 1.5 M to about 4.5 M, but not in salt concentrations below 1.0 M (pg. 501, left col., 1st. para., lines 1-16; pg. 503, rt. col., lines 3-6; and Figures 6-7, with the description thereof at pg. 504, first full para.).

Thus, given the above-noted claim dependencies, one of ordinary skill in the art would fully understand what is meant by a protein activity that is "salt tolerant."

It is also the position of the PTO that claims 1-5, 11-15, 64, and 65 are indefinite for the recitation of "low affinity." Applicants respectfully disagree. The fact that claim language, including terms of degree, may not be precise, does not automatically render the claim indefinite under 35 U.S.C. 112, second paragraph. *Seattle Box Co., v. Industrial Crating & Packing, Inc.,* 731 F.2d. 818, USPQ 568 (Fed. Cir. 1984). Acceptability of the claim language depends on whether one of ordinary skill in the art would understand what is claimed, in light of the specification. *See Manual of Patent Examining Procedure* 2173.05(b).

Applicants submit that the specification clearly defines a protein that binds oxygen with a low affinity as an oxygen-binding protein that has ability to unload oxygen into tissue when low oxygen conditions exist in the tissue (see pg. 2, lines 7-24). Furthermore, because the oxygen affinity for various types of globins were known before the present application was filed (*see, e.g.*, Gilles-Gonzales et al., "Heme-Based Sensors, Exemplified by the Kinase FixL, Are a New Class of Heme Protein with Distinctive Ligand Binding and Autoxidation," *Biochemistry* 33:8067-8073, Table 2 at pg. 8071 and pg. 8070, rt. col., 1st full para.) (1994) (attached hereto as Exhibit 4)), and "low affinity" is clearly a term of art known to those skilled in the art (Id.), applicants submit that one of ordinary skill in the art, having read the present application, would understand the meaning of the limitation "low affinity."

For all the foregoing reasons, the rejection of claims 1-5, 11-15, 50, 53, 64, and 65 under 35 U.S.C. § 112 (2nd para.) for indefiniteness is improper and should be withdrawn.

The rejection of claims 1-6 under 35 U.S.C. § 102(b) as anticipated by Zhang et al., "Signal Transduction in the Archaeon *Halobacterium salinarium* is Processed Through Three Subfamilies of 13 Soluble and Membrane-Bound Transducer Proteins," *Proc. Natl. Acad. Sci. USA* 93:4649-4654 (1996) ("Zhang"), along with the sequence alignment attached to the office action Genbank accession T44978, is respectfully traversed.

Zhang relates to the identification of signal transducer proteins from *H. salinarium*. The PTO has taken the position that Zhang teaches a heme binding protein, HtB, that is identical to SEQ ID NO:2 of the present invention. Applicants respectfully disagree.

First and foremost, the full protein sequence referred to by the PTO is not disclosed in Zhang but instead is disclosed in GenBank Accession T44978, which became publicly available on January 21, 2000 (at the time a sequence revision was made). Because the publication date of GenBank accession T44978 is after the filing date of the present application, GenBank accession T44978 is not available as prior art. Secondly, Zhang only reports partial amino acid sequences of the HtB protein (*see* Figure 3), with the fragment shown in Figure 3A possessing a highly conserved signaling domain. One of ordinary skill in the art is left to speculate as to the full amino acid sequence of the HtB protein identified in Zhang, because the full nucleotide sequence is likewise not disclosed. Finally, Zhang does not teach the isolation of the heme binding protein, but instead only shows the distribution of the protein in halobacterial cells using radiolabeled cell fractions to identify the protein in the soluble and membrane fractions (*see* pg. 4653, Fig. 5 and col. 1, lines 15-23). The protein itself was not purified from the cell fractions and, hence, cannot be considered isolated. Thus, Zhang does not teach an "isolated bacterial heme binding protein" as presently claimed.

At a minimum, the PTO should acknowledge that Zhang does not teach or suggest the amino acid sequence of SEQ ID NO: 2. Therefore, claim 6 should be allowed over Zhang.

For all these reasons, Zhang cannot anticipate the present invention and the rejection of claims 1-6 under 35 U.S.C. § 102(b) as anticipated by Zhang should be withdrawn.

The rejection of claims 1, 3-5, 11, 13-15, and 64 under 35 U.S.C. § 102(b) as anticipated by U.S. Patent No. 5,635,207 to Grinstaff et al. ("Grinstaff") is respectfully traversed.

Grinstaff relates to a composition for the *in vivo* delivery of a biologic, where the biologic is associated with either (i) a polymeric shell formulated from a biocompatible material or (ii) a dispersing agent encased within the shell. Grinstaff teaches that polymeric shells prepared from hemoglobin have a high-oxygen binding capability and are useful as blood substitutes.

In contrast, the presently claimed invention is drawn to an "isolated bacterial heme binding protein" as recited in claim 1; a blood substitute that includes "a bacterial heme binding protein" as recited in claim 11; and a fragment of the isolated protein of claim 1 as recited in claim 64.

Claims 1 and 11 (as well as claims dependent thereon) are composition claims; they are not process claims. Moreover, contrary to the assertion by the PTO, the source of the protein as "bacterial" is a property of the claimed protein rather than a process limitation. Because the protein does *not* exist in isolated form in nature, designation of its source organism defines a property that can distinguish one heme binding protein from another (e.g., bacterial from mammalian). Therefore, the source limitation "bacterial" which appears in claims 1 and 11 (and thus claims dependent thereon) must be given patentable weight; it cannot be disregarded as a mere process limitation.

Because Grinstaff does not teach or suggest an "isolated bacterial heme binding protein" as recited in claim 1 or a blood substitute that includes "a bacterial heme binding protein" as recited in claim 11, Grinstaff cannot anticipate the claimed invention.

For this reason, the rejection of claims 1, 3-5, 11, 13-15, and 64 under 35 U.S.C. § 102(b) as anticipated by Grinstaff should be withdrawn.

The rejection of claims 1, 3-5, 11, 13-15, and 64 are rejected under 35 U.S.C. § 102(b) as anticipated by Sugimoto et al., "Myoglobin Mutants Giving the Largest Geminate Yield in CO Rebinding in the Nanosecond Time Domain," *Biophysical Journal* 75:2188-2194 (1998) ("Sugimoto").

Sugimoto relates to the testing of the capacity of carbon monoxide rebinding by recombinant mutant sperm whale myoglobin proteins isolated and purified using an *E. coli* expression system. It is the position of the PTO that Sugimoto anticipates the heme-binding protein of the present invention because it teaches a recombinant myoglobin isolated from a

bacterial expression system. Even though the mutant proteins of Sugimoto are isolated from a recombinant *E. coli*, the genetic origin of the protein is not the recombinant host but instead a sperm whale, which is a mammalian source rather than a bacterial source. As noted above, the source from which the protein is isolated does identify a property that can distinguish one heme binding protein from another (e.g., bacterial from mammalian). Because Sugimoto does not teach or suggest an "isolated bacterial heme binding protein" as recited in claim 1 or a blood substitute that includes "a bacterial heme binding protein" as recited in claim 11, Sugimoto cannot anticipate the claimed invention.

Thus, the rejection of claims 1, 3-5, 11, 13-15, and 64 as anticipated by Sugimoto is improper and should be withdrawn.

The rejection of claims 1, 3-5, 11, 13-15, and 64 under 35 U.S.C. § 102(b) as anticipated by Zhao et al., "A Double Mutant of Sperm Whale Myoglobin Mimic the Structure and Function of Elephant Myoglobin," *J. Biol. Chem.* 270(35):20763-20774 (1995) ("Zhao") is respectfully traversed.

Zhao relates to a L29F/H64Q double mutant of sperm whale myoglobin having a low O_2 binding affinity, where the mutant proteins are recombinant sperm whale proteins obtained by using a bacterial expression system. Thus, the proteins of Zhao are mammalian proteins, not bacterial proteins, because the genetic origin of the mutant proteins is not the recombinant host but instead a sperm whale. Because Zhao does not teach or suggest an "isolated bacterial heme binding protein" as recited in claim 1 or a blood substitute that includes "a bacterial heme binding protein" as recited in claim 11, Zhao cannot anticipate the claimed invention.

Thus, the rejection of claims 1, 3-5, 11, 13-15, and 64 under 35 U.S.C. § 102(b) as anticipated by Zhao is improper and should be withdrawn.

The rejection of claims 1-5, 11-15, and 64-65 under 35 U.S.C. 102(b) as anticipated by Gong et al., "Structure of a Biological Oxygen Sensor: A New Mechanism, for Heme-Driven Signal Transduction," *Proc. Natl. Acad. Sci. USA* 95:15177-15182 (1998) ("Gong") is respectfully traversed in view of the above amendments.

Gong teaches a FixL protein that is isolated from the bacterium *Bradyrhizobium japonicum* and has both a heme binding domain and a histidine kinase signal transducing domain.

Applicants submit that the FixL protein of Gong does not anticipate the claimed invention because it does not possess a heme binding domain that has "at least 20% identity to a myoglobin heme binding domain having an amino acid sequence of SEQ ID NO: 76" as recited in claims 1 and 11. As evidence of this, the undersigned attorney performed a CLUSTALW sequence alignment (attached hereto as Exhibit 5) comparing the amino acid sequence of the FixL protein sequence of Gong (Genbank Accession P23222, included with Exhibit 5) with the sperm whale myoglobin protein sequence of SEQ ID NO: 76. The CLUSTALW analysis shows that there is less than 20% identity between FixL and the polypeptide of SEQ ID NO: 76. Because Gong fails to teach an isolated bacterial heme-binding protein that possesses the recited property, Gong cannot anticipate the invention of claims 1 and 11 let alone claims dependent thereon.

Thus, the rejection of claims 1-5, 11-15, and 64-65 under 35 U.S.C. 102(b) as anticipated by Gong is improper and should be withdrawn.

Applicants submit that the objection to claims 12-16, 49-54, and 64-65 is overcome by the above amendments to these claims.

In view of all of the foregoing, applicants submit that this case is in condition for allowance and such allowance is earnestly solicited.

Respectfully submitted,

Date: August 5, 2003

Edwin V. Merkel
Registration No. 40,087

NIXON PEABODY LLP
Clinton Square, P.O. Box 31051
Rochester, New York 14603-1051
Telephone: (585) 263-1128
Facsimile: (585) 263-1600

Certificate of Mailing - 37 CFR 1.8(a)

I hereby certify that this correspondence is being deposited with the United States Postal Service as first class mail in an envelope addressed to: Commissioner for Patents P.O. Box 1450 Alexandria, VA 22313-1450, on the date below.

8/5/03 Jane C. Wirszyla

Date Jane C. Wirszyla

Salt D pendent Stability and Unfolding of [F 2-S2] Ferredoxin of *Halobacterium salinarum*: Spectroscopic Inv stigations

Amal K. Bandyopadhyay and Haripalsingh M. Sonawat
Department of Chemical Sciences, Tata Institute of Fundamental Research, Mumbai 400 005, India

ABSTRACT Ferredoxin from the haloarchaeon *Halobacterium salinarum* is a 14.6-kDa protein with a [Fe2-S2] center and is involved in the oxidative decarboxylation of 2-oxoacids. It possesses a high molar excess of acidic amino acid residues and is stable at high salt concentration. We have purified the protein from this extreme haloarchaeon and investigated its salt-dependent stability by circular dichroism, fluorescence, and absorption techniques. The predominantly β-sheeted protein is stable in salt concentrations of $\geq$1.5 M NaCl. At lower concentrations a time-dependent increase in fluorescence intensity ratio (I_{360}:I_{330}), a decrease in the absorption at 420 nm, and a decrease in ellipticity values are observed. The rate of fluorescence intensity change at any low salt concentration is the highest, followed by absorption and ellipticity. This suggests that at low salt the unfolding of ferredoxin starts with the loss of tertiary structure, which leads to the disruption of the [Fe2-S2] center, resulting in the loss of secondary structural elements.

INTRODUCTION

Archaea are a group of ancient organisms that are found in adverse environmental conditions like high salinity, high temperature, and high or low pH conditions; the haloarchaea grow in nearly NaCl-saturated concentration (Ginzburg et al., 1970; Kushner, 1978). This group of organisms is known as "extreme halophiles." The intracellular salt concentration is found to be very high in extreme halophiles (Ginzburg et al., 1970). As a consequence, universally conserved molecules like proteins and nucleic acids may have also adapted to function at such a high salt concentration (Eisenberg et al., 1992). Indeed, several of the functional proteins in haloarchaea have been reported to be halo-adapted in that they are structurally intact and functionally active at high salt concentrations only (Eisenberg et al., 1992; Lanyi, 1974; Hecht et al., 1990; Krishnan and Altekar, 1993; Madan and Sonawat, 1996). For biochemical characterization of the effect of salt on these proteins and enzymes, it is necessary to purify them to homogeneity. In the recent past halophilic proteins have been amenable to purification by ammonium sulfate-mediated chromatography (Mevarech et al., 1976; Von der Haar, 1976; Leicht et al., 1978; Werber and Mevarech, 1978; Danson et al., 1984; Bonete et al., 1986; Guinet et al., 1988).

Ferredoxin (Fd) is a ubiquitous protein found in almost all living organisms, with variations in molecular weight, composition of amino acids, and primary sequence. Because of its structural diversity and presence in a wide range of phylogenetically unrelated organisms, from primitive anaerobic prokarya to all eukarya, this protein has served as a useful molecular probe for understanding biological evolution. The nonheme, iron-sulfur center in ferredoxins acts as an electron carrier in a variety of biochemical processes like carbon metabolism, nitrogen fixation, photosynthesis, and steroid hydroxylation (Hall and Evans, 1969; Buchanan and Arnon, 1970). Ferredoxin from *Halobacterium salinarum*, like that from spinach, contains a [Fe2-S2] center, unlike in the case of *Escherichia coli*, and its function as an electron carrier in the decarboxylation of α-ketoacids has been demonstrated (Kerscher and Oesterhelt, 1977). The crystal structure of ferredoxin of *Haloarcula marismortui* has been reported (Frolow et al., 1996).

Because the halophilic proteins, in general, need high salt concentrations to maintain their biological activity, it is pertinent to study the possible role of salt in the stability of its structure. Very few haloarchaeal proteins, such as glutamate and malate dehydrogenases from *Haloarcula marismortui* (Leicht et al., 1978; Mevarech et al., 1977; Madern and Zaccai, 1997; Ebel et al., 1999) and *Haloferax mediterranei* (Ferrer et al., 1998), glucose dehydrogenase from *H. salinarum* (Madan and Sonawat, 1996) and *Haloferax mediterranei* (Bonete et al., 1996), and dihydrolipoamide dehydrogenase (Jolley et al., 1997) and dihydrofolate reductase from *H. volcanii* (Zusman et al., 1989; Pieper et al., 1998), have been examined in detail for the role of salt in the stabilization of protein structure. The paucity of data from such studies may stem from the fact that halophilic proteins need special approaches for their purification and characterization. We report here a procedure for purifying the ferredoxin from *H. salinarum* (HsFd) that not only gives a higher yield but also a preparation with the maximum A_{420}/A_{275} ratio that characterizes the integrity of the protein. We provide evidence that this protein requires a high salt concentration to remain intact over extended periods of time, and withdrawal of salt results in the gradual loss of its tertiary and secondary structures, ultimately leading to its denaturation. Our data conclusively show that *H. salinarum*

Received for publication 8 December 1999 and in final form 29 February 2000.

Address reprint requests to Dr. Haripalsingh M. Sonawat, Department of Chemical Sciences, Tata Institute of Fundamental Research, Homi Bhabha Road, Mumbai 400 005, India. Tel.: 91-22-215-2971; Fax: 91-22-215-2110; E-mail: hms@tifr.res.in.

0006-3495/00/07/501/10 $2.00

absorption coefficient at 420 nm. These two methods gave similar estimates, and typically 3 mg purified HsFd per liter of culture was obtained. The protein was highly stable when stored at 25°C in 4.5 M NaCl at pH 7.3 and did not show changes in its optical spectra or A_{420}:A_{275} ratio over extended periods of time (>120 h).

Estimate of secondary structure contents by CD spectra

The far-UV CD spectra of HsFd in high and low NaCl are presented in Fig. 1. In high NaCl, a double minimum was observed near 210 and 217 nm. However, a completely different spectrum was seen for the protein in low NaCl. Intermediate NaCl concentrations showed a gradual decrease in the ellipticity. Ellipticity at 217 nm has been plotted as a function of time at different NaCl concentrations (Fig. 2). In concentrations above 1.0 M NaCl, ellipticity remains more or less unchanged with the progress of time, whereas at lower concentrations the protein shows time-dependent reduction in ellipticity values, indicating the loss of secondary structure. This loss of secondary structure became prominent at concentrations below 0.5 M NaCl. The CD spectra for several such NaCl concentrations were analyzed for secondary structural elements by the method of Andrade et al. (1993). The results for the high- and low-salt concentrations are presented in Table 1. It is seen that in high salt the protein has a predominantly β-sheet content. As the NaCl concentration is lowered, the protein loses a substantial amount of secondary structure, in which the loss of β-sheet is more pronounced than the loss of helix. The rate constants for the decrease in mean molar ellipticity at 217 nm calculated from these spectra are shown in Table 2.

Fig. 3 shows the near-UV CD spectra of HsFd in high and low NaCl. In the presence of high salt, negative peaks were observed at 267, 286, and 293 nm. These minima are indicative of specific tertiary connectivities with the aromatic side chains (Krishnan and Altekar, 1993). At low NaCl concentration, however, these ellipticity values were lower, indicating the loss of tertiary structure.

In the visible region, the CD spectra (Fig. 3) of ferredoxin in high NaCl showed positive peaks at 362 and 429 nm and minima at 328, 339, 397, 510, and 553 nm. In low NaCl an overall decrease in the ellipticity was again observed.

Secondary structure in the vicinity of the [Fe2-S2] center changes in low salt

In Fig. 4, the absorbance spectra of HsFd for high and low NaCl concentrations are presented. Ferredoxin shows prominent absorption peaks at 277, 329, 421, and 467 nm in high salt. These are characteristic of halophilic ferredoxin (Werber and Mevarech, 1978). At lower salt concentration specific enhancement at 277 nm was observed, whereas other major peaks became relatively less intense. We have monitored the characteristic ratio A_{420}:A_{275} for various salt concentrations as a function of time (Fig. 5). The data show that the ratio does not change for a long time if the salt concentrations are in the range of 2.0–4.5 M NaCl, but at lower salt concentrations a reduction in this ratio is clearly seen. We conclude from these data that the [Fe2-S2] center is destabilized upon exposure of HsFd to low salt.

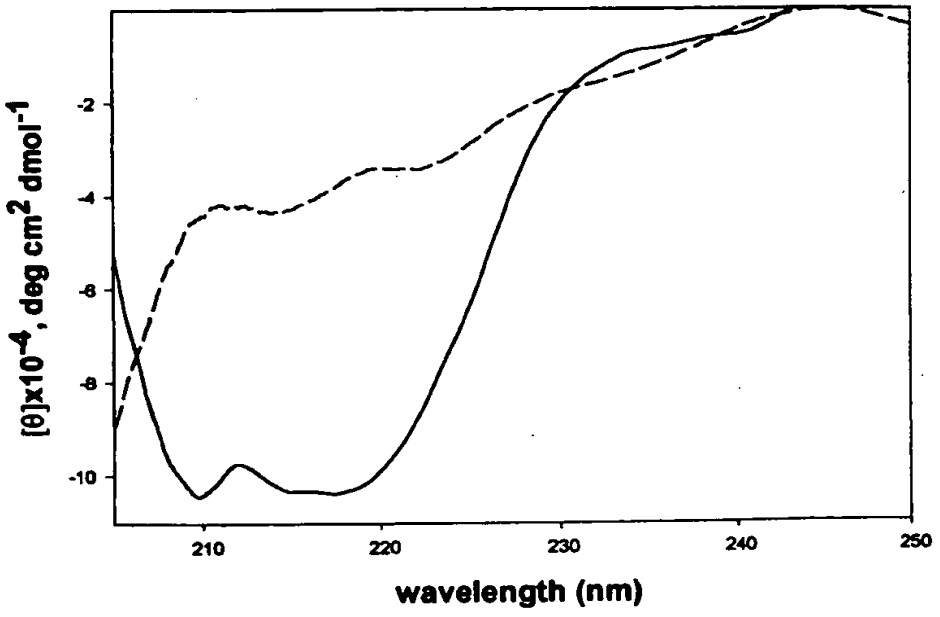


FIGURE 1 Far-UV CD spectra of ferredoxin of *Halobacterium salinarum*. ——, Salt concentration of 4.5 M NaCl; – – –, after incubation in 0.1 M NaCl for 60 h.

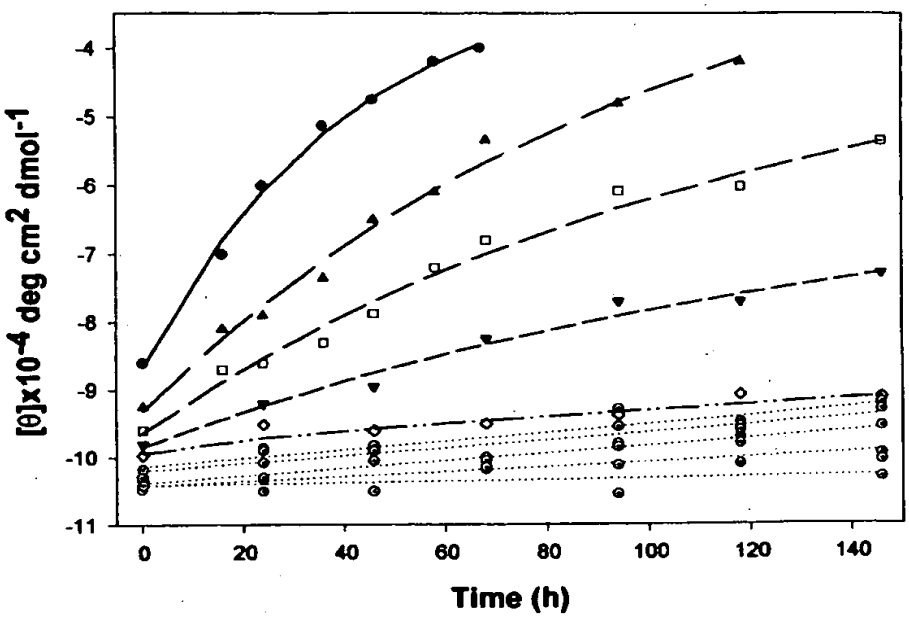


FIGURE 2 Kinetics of unfolding of ferredoxin of *H. salinarum*. CD ellipticity at 217 nm with protein in 0.1 M (●), 0.25 M (▲), 0.5 M (□), 1.0 M ([itrif), 1.5 M (◇), and 2.0–5.0 M (○) NaCl. The symbols represent experimental points and the lines correspond to two-state exponential fits.

Transfer of HsFd to low salt changes fluorescence characteristics of tryptophan residues

Fig. 6 shows the fluorescence spectra of ferredoxin at high and low NaCl. An emission maximum at 335 nm was observed for ferredoxin at high salt, whereas at low salt concentration the emission maxima shifted to 355 nm. All of the intermediate NaCl concentrations showed values intermediate between these two extremes and therefore were clearly salt concentration dependent. We have monitored the I_{360}:I_{330} ratio of the protein at different NaCl concentrations over a period of several days (Fig. 7). In high concentrations of NaCl the ratio remains more or less invariant, whereas at concentrations below 1.0 M a time-dependent increase was observed. The redshift of the λ_{em} and an increase in the I_{360}:I_{330} ratio indicate exposure of the tryptophan residues to an aqueous environment. Unfolding of the protein at low salt concentration is clearly apparent. An estimate of the rate constants of this unfolding process at the various salt concentrations is presented in Table 2. The data show that the rate of unfolding is dependent on the salt concentration. The HsFd unfolding rate is enhanced when it is transferred from high to low salt.

TABLE 1 Secondary structure elements of *H. salinarum* ferredoxin calculated from CD spectra in 10 mM phosphate buffer (pH 7.3) with the indicated salt concentrations

NaCl concentration (M)	α-Helix	β-Sheet	Others
4.5	23.0 ± 3	43.0 ± 2	32.0 ± 3
0.1	10.3 ± 4	15.5 ± 2	74.2 ± 4
% change	55.0	64.0	131

The fractions were determined at ambient temperature by the method of Andrade *et al* (1993).

DISCUSSION

H. salinarum is a facultative anaerobe that belongs to the archaeal domain. It grows in high salt environments. This microbe has adapted the mechanism of osmoprotection by accumulating a high concentration of salt inside the cell (Ginzburg et al., 1970). As a result of this, proteins from haloarchaea, in general, require high salt to maintain their structural and functional integrity. The structures of a few of the enzymes like malate dehydrogenase, glutamate dehydrogenase, dihydrofolate reductase, lactate dehydrogenase, and glucose dehydrogenase have been studied in great detail

TABLE 2 Rate constants (k) for salt-dependent unfolding of ferredoxin from *Halobacterium salinarum* measured by CD, fluorescence, and optical absorbance

Salt concentration (M)	$10^{-3} \times k$ (h^{-1})		
	CD	Fluorescence	Absorbance
0.10	24.0 ± 0.003	62.0 ± 0.003	23.0 ± 0.002
0.25	9.6 ± 0.002	20.0 ± 0.002	12.0 ± 0.003
0.50	8.6 ± 0.002	15.0 ± 0.005	10.0 ± 0.003
1.00	7.0 ± 0.003	10.0 ± 0.004	8.6 ± 0.004
1.50	4.8 ± 0.010	5.9 ± 0.070	5.5 ± 0.051

The rate constants for CD experiments were determined from ellipticity measurements at 217 nm, for fluorescence from the I_{360}:I_{330} ratio, and for absorbance from the A_{420}:A_{275} ratio. The protein was in 10 mM phosphate buffer (pH 7.3), and measurements were performed at ambient temperature.

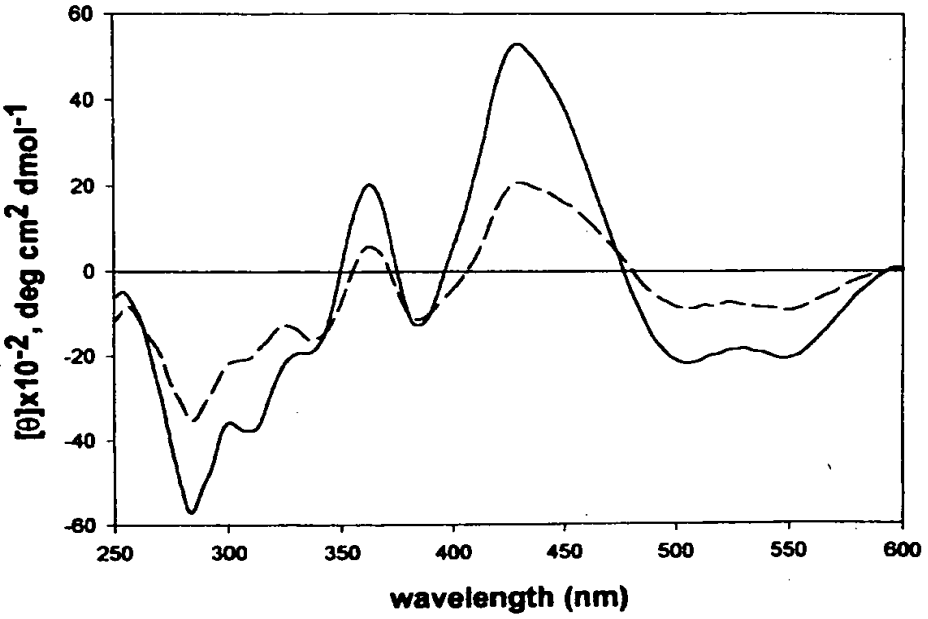


FIGURE 3 UV-visible CD spectra of ferredoxin of *H. salinarum* in 4.5 M NaCl (——) and after incubation in 0.1 M NaCl for 60 h (– – –).

as a function of salt concentration (Eisenberg et al., 1992; Hecht et al., 1990; Madan and Sonawat, 1996), and it has been shown that at low salt these enzymes lose their activities. The effect of salt on the structure of some halophilic proteins has been explored in great detail (Zaccai et al., 1989; Cendrin et al., 1993; Bonnete et al., 1994; Madern and Zaccai, 1997; Ebel et al., 1999). Haloarchaeal ferredoxin acts as a cofactor in decarboxylation reaction catalyzed by oxidoreductase. Therefore, the study of the effect of salt on its structure as well as its function may provide useful information about the structural biology of this protein.

Ferredoxin was purified earlier by a procedure that involved its exposure to low salt (Kerscher and Oesterhelt,

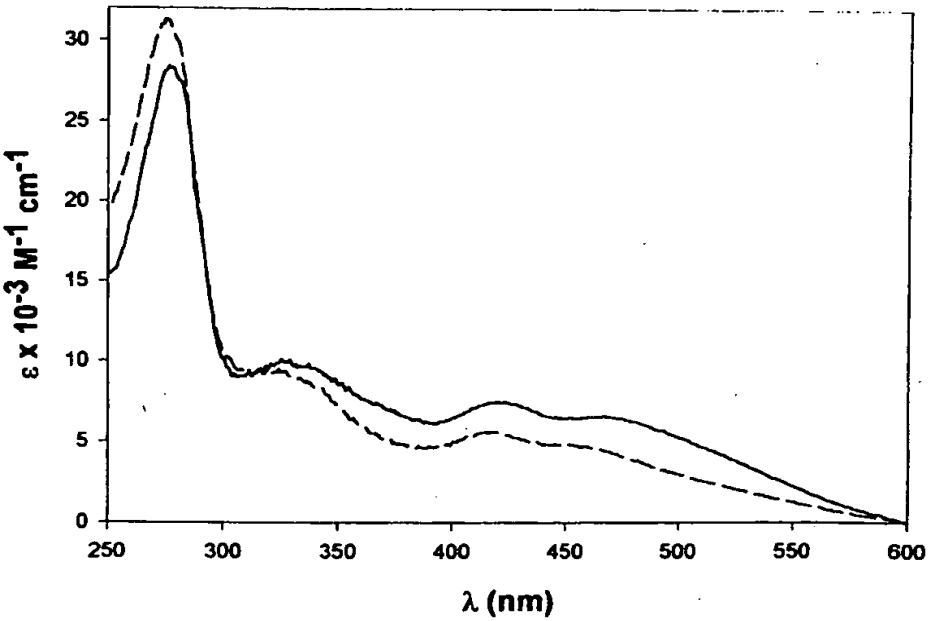


FIGURE 4 Absorption spectra of *H. salinarum* ferredoxin in 4.5 M NaCl (——) and after incubation in 0.1 M NaCl for 60 h (– – –).

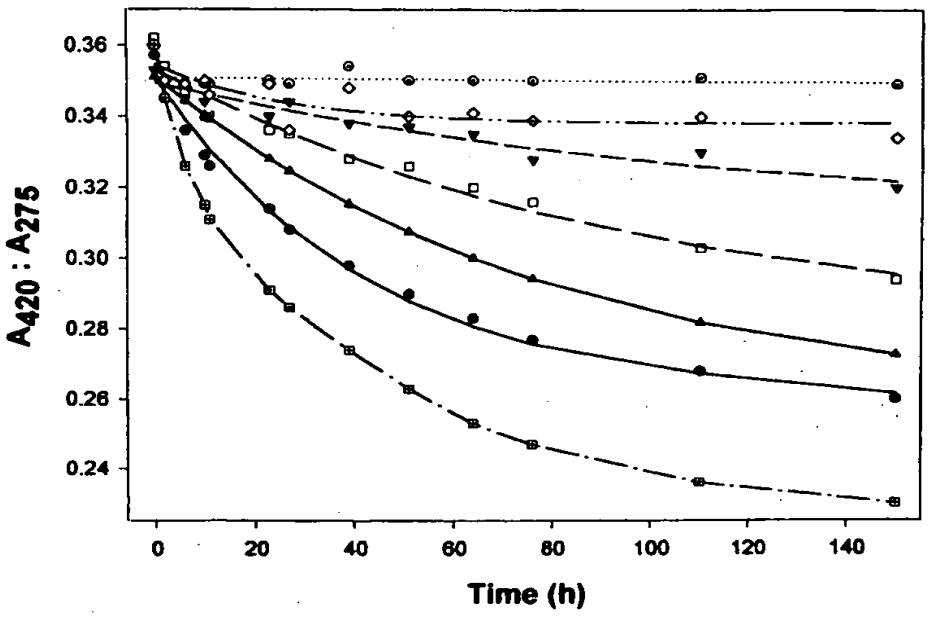


FIGURE 5 Unfolding kinetics of *H. salinarum* ferredoxin. The absorption ratio A_{420}:A_{275} corresponds to the unfolding in 0.05 M NaCl (⊠) and 4.5 M NaCl (┄┄). Other symbols are as in Fig. 2. The symbols represent experimental points, and the lines correspond to two-state exponential fits.

1976). This was improved by Werber and Mevarech (1978), who employed a method that maintains a high salt concentration throughout the purification procedure. In view of the instability of the *H. salinarum* ferredoxin at low salt, we used a procedure retaining several features of the latter method. In our procedure, a yield of 3 mg/liter cell culture with a high A_{420}:A_{275} ratio of 0.35 has been reproducibly obtained over several batches of purification. This may be contrasted with the nonhalophilic procedure, which has a lower yield (0.4 mg/liter cell culture) and an A_{420}:A_{275} ratio of 0.25 (Kerscher et al., 1976), and the modified procedure (Werber and Mevarech, 1978). The introduction of a Sepharose-4B column results in additional purification. This has also been reported for ferredoxin from *H. marismortui* (Werber and Mevarech, 1978). The most significant point in the procedure reported here is that the salt concentration has

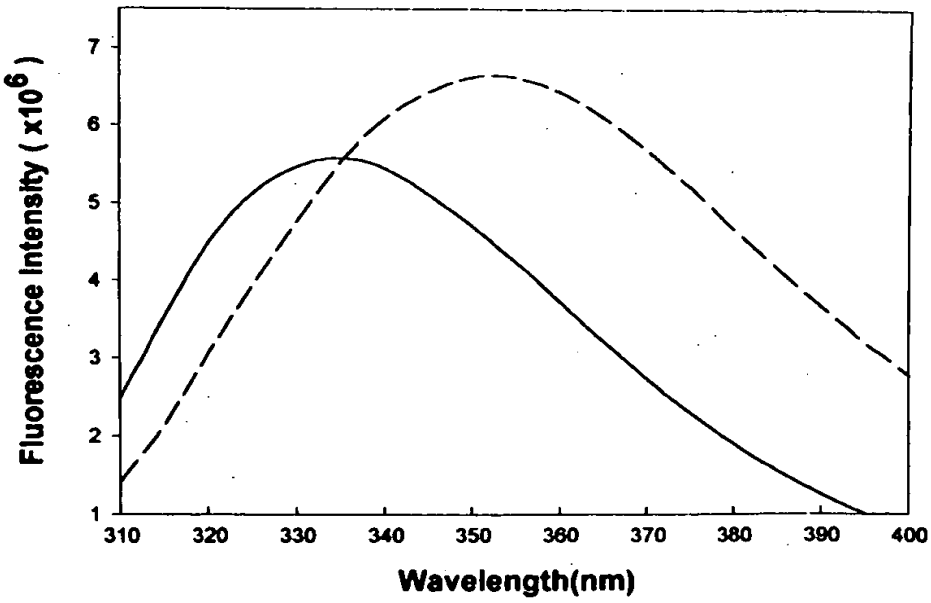


FIGURE 6 Fluorescence emission spectra of *H. salinarum* ferredoxin in 10 mM phosphate buffer containing 4.5 M NaCl (——) and 60 h after incubation in the same buffer with 0.1 M NaCl (– – –).

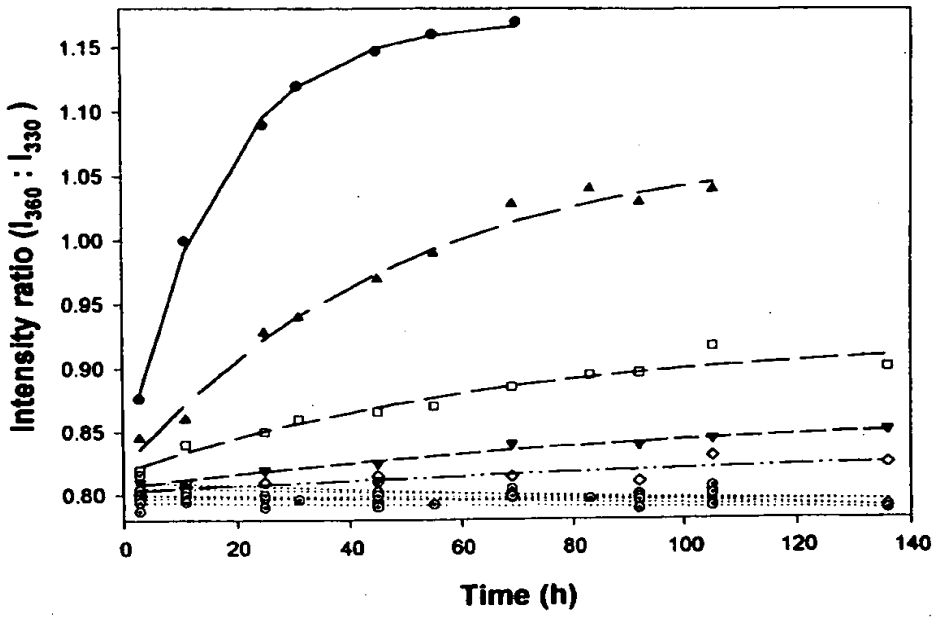


FIGURE 7 Unfolding kinetics of *H. salinarum* ferredoxin from the fluorescence intensity ratio I_{360}:I_{330}. The protein is in 10 mM phosphate buffer (pH 7.3) containing the indicated salt concentrations. Symbols are as in Fig. 2. The symbols are experimental points and the lines are the best fits, assuming a two-state process.

not been reduced below 1 M in any of the steps of purification. All of the purification steps were carried out in buffer containing ammonium sulfate, which is known to stabilize halophilic proteins (Eisenberg et al., 1992). Thus, during the purification procedure, ferredoxin could be kept intact in the absence of sodium chloride. Exposure to low salt has deleterious effects on halophilic proteins, which might be the prime factor for the observation of a lower absorption ratio in the earlier procedures. As we show later, exposure of purified HsFd to low salt may reduce the A_{420}:A_{275} ratio to <0.25 in ~48 h. Prolonged exposure to low salt leads to aggregation and eventual protein precipitation. These features explain why the purified HsFd had a lower A_{420}:A_{275} ratio in apparently electrophoretically homogeneous preparations obtained in earlier procedures. Although, ammonium sulfate-mediated chromatography has advantages over the nonhalophilic procedure, it has to be ensured that the protein is not exposed to ammonium sulfate for extended periods, because in such cases protein precipitation occurs and may result in a low yield.

HsFd has two tryptophan residues that are localized in two distinct regions of the protein, namely, W16 at the N-terminal end and W59 in the vicinity of the conserved [Fe2-S2] center. Our observation of a fluorescence emission maximum of 335 nm implies that these tryptophans are in a hydrophobic environment. At low salt, an increase in quantum yield and a red shift as high as 20 nm in its emission maximum suggest that local environments of tryptophans are affected, leading to their transfer to more aqueous environment. This may reflect the loss of secondary and tertiary structure (Chen et al., 1967; Altekar, 1977a,b). This unfolding process is kinetically slow and continues for several hours, even days, to reach an apparent end point by spectroscopic criteria. The fluorescence intensity increase upon exposure to low salt has also been seen in other cases (Irace et al., 1981; Hargrove et al., 1994). The increased intensity may result when the tryptophans are relieved from an internal quenching in the folded state. Overall, the red-shift, along with the change in the spectral behavior, could be ascribed to loss of tertiary and/or secondary structure of ferredoxin at low salt (Staniforth et al., 1998).

The CD of salt-stabilized HsFd indicates 1) predominant β-sheet structure (~43%), 2) low α-helical content (23%), and 3) an intact [Fe2-S2] center. Our observation of a low α-helical content in the case of HsFd is in agreement with earlier reports (Werber and Mevarech, 1978; Frolow et al., 1996). The crystal structure of *Haloarcula marismortui* ferredoxin has been reported to have a β-sheet content of 25% (Frolow et al., 1996). Exposure to low salt induces unfolding, as indicated by an incremental random structure and a decrease in the secondary structure elements, in which the loss of β-sheet is greater than the loss of α-helix. In addition, low salt also affected the tertiary connectivities of the aromatic residues and the [Fe2-S2] center.

That the HsFd structure is destabilized in low salt is again revealed by optical spectra. The absorption spectrum for the haloarchaeal ferredoxin in its salt-stabilized state exhibits maxima at 275 and 420 nm. The A_{420}:A_{275} ratios for various ferredoxins have been determined and are found to have a characteristic value for any given species. This ratio can therefore be taken as a signature of the integrity of the [Fe2-S2] center and that of functionally intact protein.

Lower values of this ratio, if they occur because of external perturbation, may therefore be ascribed to unfolding (Keresztes-Nagy and Margoliash, 1966; Werber and Mevarech, 1978). Our observation that this ratio gradually decreases upon exposure to low salt in a time-dependent manner clearly suggests that the integrity of the [Fe2-S2] center is affected by protein unfolding.

The time-dependent destabilization of tertiary and secondary structure and the integrity of the [Fe2-S2] center were followed at salt concentrations in the range 0.1–5 M NaCl. A comparison of the rate constants (Table 2) reveals that for each of the salt concentrations the change is at maximum as determined by fluorescence and at minimum by CD spectroscopy. However, given that these disparate techniques monitor different aspects of overall protein structure and these three independent methods show a similar trend in the manner in which HsFd unfolding occurs, the half-lives determined by these methods are in general agreement with one another (Fig. 8). At the lowest salt concentration studied, ferredoxin tends to aggregate and precipitate after ~70 h.

Three distinct phases of structural destabilization are observed in our experiments. The low-salt range (0.1–0.25 M) corresponds to electrostatic charge repulsion. This originates largely from the large excess of negatively charged residues, aspartates and glutamates. HsFd has a total of 21 each of Asx and Glx (Kerscher et al., 1976), as compared to 10 Asx and 14 Glx of *E. coli* (Knoell and Knappe, 1974). Negatively charged residues cover the entire surface, barring the [Fe2-S2]center of *H. marismortui* ferredoxin (Frolow et al., 1996), which shows 88% homology to HsFd. These residues exist preferentially on the solvent-exposed side of α-helices and in loops that connect the various secondary structure elements. The crystal structure of this protein also suggests that the negative charges are shielded from each other by extensive solvation and are thus noninteracting (Frolow et al., 1996). The stability of HsFd in this range of salt concentration is, therefore, due to screening of these charges and the consequent stabilization of the α-helices and the loops. This is even more evidently the case for the amphipathic helix in the N-terminal 22-residue extra segment reported in *H. marismortui* ferredoxin. Because of the high homology, this helix is also expected in the HsFd, where the N-terminal segment contains nine negatively charged residues and no positive charges. This hyperacidic region of the protein, it is reasonable to assume, contributes largely to the stability of HsFd in the low-salt concentration range. Enhanced stability at lower pH values, where the side-chain carboxyl groups of these residues are undissociated (protonated), and in the presence of lower concentration of di- and multivalent cations supports the electrostatic charge screening (Baxter, 1959; Brown, 1963, 1964a,b, 1965; Lanyi, 1974). In addition, the negatively charged residues have an excellent solvation capacity, and the binding of water and/or salt to halophilic proteins is ~10 times more than that for nonhalophilic proteins (Ebel et al., 1992, 1995; Bonnete et al., 1993). Indeed, the surface water in *H. marismortui* ferredoxin has been shown to have 40% more

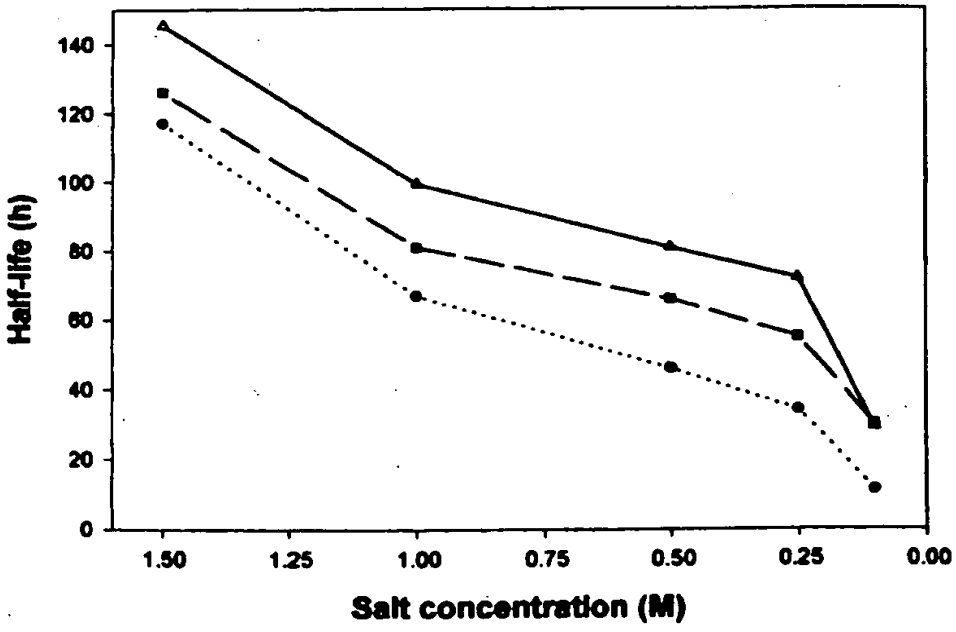


FIGURE 8 Half-life of *H. salinarum* ferredoxin determined by fluorescence (○), absorption (□), and CD (△) as a function of NaCl concentration.

hydrogen bonds (Frolow et al., 1996). The stabilization of the hyperacidic segment in HsFd is probably due to excessive solvation. The charge screening seems to continue in the intermediate salt concentration range as well, although its contribution to the overall stability is much less. In the high-salt concentration range the hydrophobic effects play a dominant role in the stabilization of halophilic proteins (Lanyi, 1974). The electrostatic interactions are affected to a smaller extent in this range (Elcock and McCammon, 1998). HsFd also possesses a large excess of valine, glycine, alanine, and tyrosine (Kerscher et al., 1976) in comparison to the nonhalophilic counterpart (Knoell and Knappe, 1974). The contacts among these residues are expected to be affected by a change in the salt concentration in this range. A comparison with the crystal structure of the homologous *H. marismortui* ferredoxin reveals that a significant fraction of these residues occur in the β-sheets. Our observations (Table 1) therefore suggest that the β-sheets are preferentially destabilized by a reduction of salt and that their stabilization contributes significantly to the stability of the protein in this concentration range. The intermediate salt range effects could be due to contributions from hydrophobic and specific residue interaction along with electrostatic charge screening. Further investigations with various cations and organic solvents with varying degrees of hydrophobicity might be helpful in evaluating the relative contributions of negative charges and hydrophobic residues to the stability of ferredoxin from this haloarchaeon.

We thank Prof. A. K. Singh, Department of Chemistry, I.I.T., Bombay, for the *H. salinarum* (M1) culture. We are grateful to Dr. L. C. Padhy and Prof. G. Govil for critical discussion and many helpful suggestions while we prepared the manuscript.

REFERENCES

Altekar, W. 1977a. Fluorescence of proteins in aqueous neutral salt solutions. I. Influence of anions. *Biopolymers.* 16:341–368.

Altekar, W. 1977b. Fluorescence of proteins in aqueous neutral salt solutions. II. Influence of monovalent cation chlorides, particularly cesium chloride. *Biopolymers.* 16:369–386.

Andrade, M. A., P. Chacon, J. J. Merelo, and F. Moran. 1993. Evaluation of secondary structure of proteins from UV circular dichroism spectra using an unsupervised learning neural network. *Protein Eng.* 6:383–390.

Baxter, R. M. 1959. An interpretation of the effects of salt on the lactic dehydrogenase of *Halobacterium salinarum*. *Can. J. Microbiol.* 5:47–57.

Bonnete, F., C. Ebel, H. Eisenberg, and G. Zaccai. 1993. A biophysical study of halophilic malate dehydrogenase in solution: revised subunit structure and solvent interactions in native and recombinant enzyme. *J. Chem. Soc. Faraday Trans.* 89:2659–2666.

Bonnete, F., D. Madern, and G. Zaccai. 1994. Stability against denaturation mechanisms in halophilic malate dehydrogenase "adapt" to solvent conditions. *J. Mol. Biol.* 244:436–447.

Bonete, M. J., M. L. Camacho, and E. Cadenas. 1986. Purification and some properties of NAD^+-dependent glutamate dehydrogenase from *Halobacterium halobium*. *Int. J. Biochem.* 18:785–789.

Bonete, M. J., C. Pire, F. I. Llorca, and M. L. Camacho. 1996. Glucose dehydrogenase from the halophilic archaeon *Haloferax mediterranei*: enzyme purification, characterization and N-terminal sequence. *FEBS Lett.* 383:227–229.

Brown, A. D. 1963. The peripheral structures of gram-negative bacteria. IV. The cation sensitive dissolution of the cell membrane of the halophilic bacterium. *Halobacterium halobium*. *Biochem. Biophys. Acta.* 75:425–435.

Brown, A. D. 1964a. Aspects of bacterial response to the ionic environment. *Bacteriol. Rev.* 28:296–329.

Brown, A. D. 1964b. The development of halophilic properties in bacterial membranes by acylation. *Biochem. Biophys. Acta.* 93:136–142.

Brown, A. D. 1965. Hydrogen ion titrations of intact and dissolved lipoprotein membranes. *J. Mol. Biol.* 12:491–508.

Buchanan, B. B., and D. I. Arnon. 1970. Ferredoxins: chemistry and function in photosynthesis, nitrogen fixation, and fermentative metabolism. *Adv. Enzymol.* 33:119–176.

Cendrin, F., J. Chroboczek, G. Zaccai, H. Eisenberg, and M. Mevarech. 1993. Cloning, sequencing and expression in *Escherichia coli* of a halophilic enzyme. *Biochemistry.* 32:4308–4313.

Chen, R. F., G. G. Vurek, and N. Alexander. 1967. Fluorescence decay times: proteins, coenzymes and other compounds in water. *Science.* 156:949–951.

Danson, M. J., R. Eisenthal, S. Hall, S. R. Kessell, and D. L. Williams. 1984. Dihydrolipoamide dehydrogenase from halophilic archaebacteria. *Biochem. J.* 218:811–818.

Ebel, C., W. Altekar, J. Langowski, C. Urbanke, E. Forest, and G. Zaccai. 1995. Solution structure of glyceraldehyde-3-phosphate dehydrogenase from *Haloarcula vallismortis*. *Biophys. Chem.* 54:219–227.

Ebel, C., P. Faou, B. Kernel, and G. Zaccai. 1999. Relative role of anions and cations in the stabilization of halophilic malate dehydrogenase. *Biochemistry.* 38:9039–9047.

Ebel, C., F. Guinet, J. Langowski, C. Urbanke, J. Gagnon, and G. Zaccai. 1992. Solution studies of the elongation factor Tu from the extreme halophilic *Halobacterium marismortui*. *J. Mol. Biol.* 223:361–371.

Eisenberg, H., M. Mevarech, and G. Zaccai. 1992. Biochemical, structural, and molecular genetic aspects of halophilism. *Adv. Protein Chem.* 43: 1–62.

Elcock, A. H., and J. A. McCammon. 1998. Electrostatic contributions to the stability of halophilic proteins. *J. Mol. Biol.* 280:731–748.

Ferrer, J., R. Cremades, C. Pire, and M. J. Bonete. 1998. Fluorescence and quenching comparative studies of halophilic and bovine glutamate dehydrogenase. *J. Photochem. Photobiol. B.* 47:148–54.

Frolow, F., M. Harel, J. L. Sussman, M. Mevarech, and M. Shoham. 1996. Insights into protein adaptation to a saturated salt environment from the crystal structure of a halophilic 2Fe-2S ferredoxin. *Nature Struct. Biol.* 3:452–458.

Ginzburg, M., L. Sachs, and B. Z. Ginzburg. 1970. Ion metabolism in a halobacterium. I. Influence of age of culture on intracellular concentrations. *J. Gen. Physiol.* 55:187–207.

Guinet, F., R. Frank, and R. Leberman. 1988. Polypeptide elongation factor Tu from *Halobacterium marismortui*. *Eur. J. Biochem.* 172:687–694.

Hall, D. O., and M. C. Evans. 1969. Iron-sulphur proteins. *Nature.* 223: 1342–1348.

Hargrove, M. S., S. Krzywda, A. J. Wilkinson, Y. Dou, M. Ikeda-Saito, and J. S. Olson. 1994. Stability of myoglobin: a model for the folding of haem proteins. *Biochemistry.* 33:11767–11775.

Hecht, K., T. Langer, A. Wrba, and R. Jaenicke. 1990. Lactate dehydrogenase from the extreme halophilic archaebacterium *Halobacterium marismortui*. *Biol. Chem. Hoppe Seyler.* 371:515–519.

Irace, G., C. Balestieri, G. Parlato, L. Servillo, and G. Colonna. 1981. Tryptophanyl fluorescence heterogeneity of apo-myoglobins. Correlation with the presence of two distinct structural domains. *Biochemistry.* 20:792–799.

Jolley, K. A., R. J. Russell, D. W. Hough, and M. J. Danson. 1997. Site-directed mutagenesis and halophilicity of dihydrolipoamide dehydrogenase from the halophilic archaeon. *Haloferax volcanii*. *Eur. J. Biochem.* 248:362–368.

Keresztes-Nagy, S., and E. Margoliash. 1966. Preparation and characterization of alfalfa ferredoxin. *J. Biol. Chem.* 241:5955–5966.

Kerscher, L., and D. Oesterhelt. 1976. A ferredoxin from halobacteria. *FEBS Lett.* 67:320–322.

Kerscher, L., and D. Oesterhelt. 1977. Ferredoxin is the coenzyme of alpha-ketoacid oxidoreductases in *Halobacterium halobium*. *FEBS Lett.* 83:197–201.

Kerscher, L., D. Oesterhelt, R. Cammack, and D. O. Hall. 1976. A new plant-type ferredoxin from Halobacteria. *Eur. J. Biochem.* 71:101–107.

Knoell, H. E., and J. Knappe. 1974. *Escherichia coli* ferredoxin, an iron-sulfur protein of the adrenodoxin type. *Eur. J. Biochem.* 50: 245–252.

Krishnan, G., and W. Altekar. 1993. Halophilic class I aldolase and glyceraldehyde-3-phosphate dehydrogenase: some salt-dependent structural features. *Biochemistry.* 32:791–798.

Kushner, D. J. 1978. Life in high salt and solute concentrations: halophilic bacteria. *In* Microbiol Life in Extreme Environments. D. J. Kushner, editor. Academic Press, London. 317–368.

Lanyi, J. K. 1974. Salt-dependent properties of proteins from extremely halophilic bacteria. *Bacteriol. Rev.* 38:272–290.

Leicht, W., M. M. Werber, and H. Eisenberg. 1978. Purification and characterization of glutamate dehydrogenase from Halobacterium of the Dead Sea. *Biochemistry.* 17:4004–4010.

Lowry, O. H., N. J. Rosebrough, A. L. Farr, and R. J. Randall. 1951. Protein measurement with Folin phenol reagent. *J. Biol. Chem.* 193: 265–275.

Madan, A., and H. M. Sonawat. 1996. Glucose dehydrogenase from *Halobacterium salinarum*: purification and salt dependent stability. *Physiol. Chem. Phys. Med. NMR.* 28:15–28.

Madern, D., and G. Zaccai. 1997. Stabilization of halophilic malate dehydrogenase from *Haloarcula marismortui* by divalent cations—effects of temperature, water isotope, cofactor and pH. *Eur. J. Biochem.* 249: 607–611.

Mevarech, M., H. Eisenberg, and E. Neumann. 1977. Malate dehydrogenase isolated from extremely halophilic bacteria of the Dead Sea. 1. Purification and molecular characterization. *Biochemistry.* 16: 3781–3785.

Mevarech, M., W. Leicht, and M. M. Werber. 1976. Hydrophobic chromatography and fractionation of enzymes from extremely halophilic bacteria using decreasing concentration gradients of ammonium sulfate. *Biochemistry.* 15:2383–2387.

Peterson, G. L. 1983. Determination of total protein. *Methods Enzymol.* 91:95–119.

Pieper, U., G. Kapadia, M. Mevarech, and O. Herzberg. 1998. Structural features of halophilicity derived from the crystal structure of dihydrofolate reductase from the Dead Sea halophilic archaeon *Haloferax volcanii*. *Structure.* 6:75–88.

Staniforth, R. A., M. G. Bigotti, F. C. C. T. Allocatelli, and M. Brunori. 1998. Unfolding of apomyoglobin from *Aplysia limacina*: the effect of salt and pH on the cooperativity of folding. *J. Mol Biol.* 275:133–148.

Von der Haar, F. 1976. Purification of proteins by fractional interfacial salting out on unsubstituted agarose gels. *Biochem. Biophys. Res. Commun.* 70:1009–1013.

Werber, M. M., and M. Mevarech. 1978. Purification and characterization of a highly acidic 2Fe-2S ferredoxin from Halobacterium of the Dead Sea. *Arch. Biochem. Biophys.* 187:447–456.

Zaccai, G., F. Cendrin, Y. Haik, N. Borochov, and H. Eisenberg. 1989. Stabilization of halophilic malate dehydrogenase. *J. Mol. Biol.* 208: 491–500.

Zusman, T., I. Rosenshine, G. Boehm, R. Jaenicke, B. Leskiw, and M. Mevarech. 1989. Dihydrofolate reductase of the extremely halophilic archaebacterium *Halobacterium volcanii*. The enzyme and its coding gene. *J. Biol. Chem.* 264:18878–18883.

Th stability of proteins in xtr me nvironm nts

Rainer Jaenicke* and Gerald Böhm†

Three complete genome sequences of thermophilic bacteria provide a wealth of information challenging current ideas concerning phylogeny and evolution, as well as the determinants of protein stability. Considering known protein structures from extremophiles, it becomes clear that no general conclusions can be drawn regarding adaptive mechanisms to extremes of physical conditions. Proteins are individuals that accumulate increments of stabilization; in thermophiles these come from charge clusters, networks of hydrogen bonds, optimization of packing and hydrophobic interactions, each in its own way. Recent examples indicate ways for the rational design of ultrastable proteins.

Addresses
*Institute of Biophysics and Physical Biochemistry, University of Regensburg, D-93040 Regensburg, Germany;
e-mail: rainer.jaenicke@biologie.uni-regensburg.de
†Institute for Biotechnology, Molecular Bioinformatics, University of Halle-Wittenberg, D-06099 Halle (Saale), Germany;
e-mail: gerald.boehm@biochemtech.uni-halle.de

Current Opinion in Structural Biology 1998, 8:738–748

http://biomednet.com/elecref/0959440X00800738

Abbreviations

CS	citrate synthase
DHFR	dihydrofolate reductase
GAPDH	glyceraldehyde 3-phosphate dehydrogenase
GluDH	glutamate dehydrogenase
Hs	*Halobacterium salinarium*
IPMDH	3-isopropylmalate dehydrogenase
LDH	lactate dehydrogenase
PGK	phosphoglycerate kinase
PRAI	phosphoribosyl anthranilate isomerase
TIM	triosephosphate isomerase
Tm	*Thermotoga maritima*

Introduction

Life on earth exhibits an enormous adaptive capacity. Except for centers of volcanic activity, the surface of our planet is 'biosphere'. In quantitative terms, the limits of the biologically relevant physical variables are –40 to +115°C (in the stratosphere and hydrothermal vents, respectively), ≤120 MPa (for hydrostatic pressures in the deep sea), $a_w \approx 0.6$ (for the activity of water in salt lakes) and ≈1<pH<11 (for acidic or alkaline biotopes). During evolution, organisms achieved viability under extreme conditions either by 'escaping' or 'compensating' the stress or by enhancing the stability of their cellular inventory. In the case of temperature and pressure, there is no alternative to mutative adaptation for survival [1]. Here, we shall review the recent progress in research on protein stabilization, focusing on thermophiles with optimum temperatures of growth of more than 60°C (for hyperthermophiles, more than 80°C) and halophiles with optimum water activities around 0.6. Studies on proteins from acidophiles and alkalophiles have been scarce. Strict barophiles have recently been isolated — thousands of microbes were isolated from the first samples collected from the Challenger Deep at ~110 MPa [2], but very few of them were truly barophilic [3•]. Their proteins are still *terra incognita*.

Limits of stability and growth

Proteins, independent of their mesophilic or extremophilic origin, consist exclusively of the 20 canonical natural amino acids. In the multicomponent system of the cytosol, these are known to undergo covalent modifications at extremes of temperature, pH and pressure (deamidation, β elimination, disulfide interchange, oxidation, Maillard reactions, hydrolysis, etc. [4]). Extremophiles must compensate for amino acid degradation either by using compatible protectants or by enhanced synthesis and repair. Little is known about the chemistry involved, for example, in the hydrothermal decomposition of proteins, and even less is known about protection and repair. Applying temperatures beyond 100°C, the thermal stabilities of the common amino acids are (Val,Leu)>Ile>Tyr>Lys>His>Met>Thr>Ser>Trp>(Asp,Glu, Arg,Cys). In many cases, the half-lives of the degradation reactions are significantly shorter than the generation time of hyperthermophilic microorganisms [5]; to this limit, biomolecules could still be resynthesized at biologically feasible rates. The temperature at which ATP hydrolysis becomes the limiting factor for viability lies between 110 and 140°C [6]. This temperature limit coincides with the temperature range at which the hydrophobic hydration of proteins vanishes and water becomes an 'ordinary solvent' [1]. Apparently, both the integrity of the natural amino acids and the formation of the hydrophobic core upon protein folding are essential for viability. Extrinsic factors and compatible solutes may enhance the stability and shift the limits of growth of prokaryotes as well as eukaryotes [7].

Fundamentals of protein stability

Proteins exhibit marginal stabilities that are equivalent to only a small number of weak intermolecular interactions [1,8]. In this respect, proteins from extremophiles do not differ strongly from their mesophilic counterparts. Their adaptation, either intrinsic or through interaction with extrinsic factors, is accompanied by only marginal increases in the free energy of stabilization. No general strategy of stabilization has yet been established. In recent years, however, well-defined increments of stability have been elucidated by analyzing ultrastable proteins and verifying their specific anomalies by rational design. As indicated by these studies, stabilization may involve all levels of the hierarchy of protein structure: local packing of the polypeptide chain, secondary and supersecondary structural elements, domains and subunits [4]. Taking thermal stability as an example, several experimental approaches have been used to assign specific structural alterations to changes in stability: selection of temperature-sensitive

mutants; systematic variations of amino acid residues in the core or in the periphery of model proteins; fragmentation of domain proteins or modifications of connecting peptides between domains; and alteration of subunit interactions by mutagenesis or solvent perturbation [1,9].

Stability refers to the maintenance of a defined functional state under extreme conditions. High-resolution structures in the crystalline state and in solution have shown that the atomic coordinates of proteins can be determined down to a resolution better than 1 Å. Even this precision, however, does not allow the calculation of the free energy of stabilization from coordinates, nor does it consider the dynamics as an essential prerequisite of protein function. The polypeptide chain may fluctuate between preferred conformations with amplitudes and angles up to 50 Å and 20°, respectively [10].

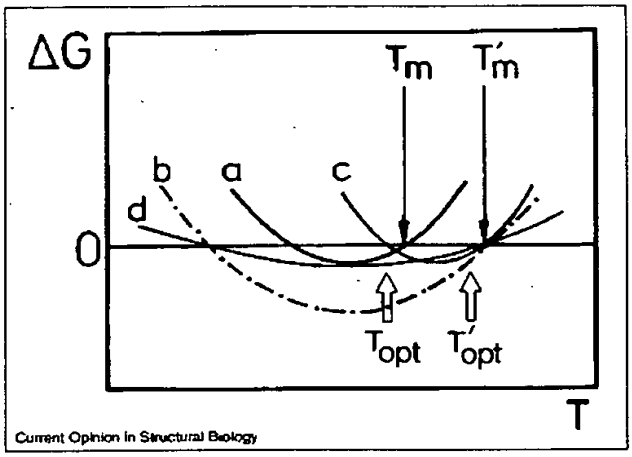
Considering extremophiles in comparison with their mesophilic counterparts, evolutionary adaptation is nothing more than the conservation of functionally important motions in such a way that, under altered physical conditions, the protein inventories of extremophiles and mesophiles are in 'corresponding states' [1]. In this context, the stability of an individual protein refers to the native state, as well as the intermediates on its pathway from the nascent or unfolded ensemble of states (U) to the functional entity (N). Evidently, in order 'to be extremophilic', a protein has to cope with the extreme conditions at all stages along its folding pathway.

Stability and folding

The driving forces that are responsible for protein folding reflect the hierarchy of contributions involved in protein stabilization, that is, on the one hand, the nearest neighbor and through-space short-range interactions that optimize packing and minimize cavity volume and, on the other hand, the entropy effects due to water release from hydrophobic surfaces [10,11]. Both the enthalpic and entropic contributions to the free energy of stabilization are affected by the extreme conditions we are dealing with. The difference in the stabilities of mesophilic and (hyper-)thermophilic proteins, $\Delta\Delta G_{N\rightarrow U}$, does not exceed ~100 kJ/mol, that is, the equivalent of a few noncovalent interactions [12•,13•]; often it is even lower, rendering the definition of general 'strategies' of thermal adaptation extremely difficult. This is especially true because the enthalpic contributions, in terms of additional stabilizing interactions, necessarily lead to the previously mentioned decrease in flexibility, which naturally corresponds to a decrease in conformational entropy. Thus, evolution had to find a balance between rigidity as a prerequisite of stability and specificity, on the one hand, and flexibility, for example, in connection with ligand interactions and degradation, on the other.

As a consequence of the parabolic temperature dependence of the free energy of stabilization, proteins exhibit

Figure 1



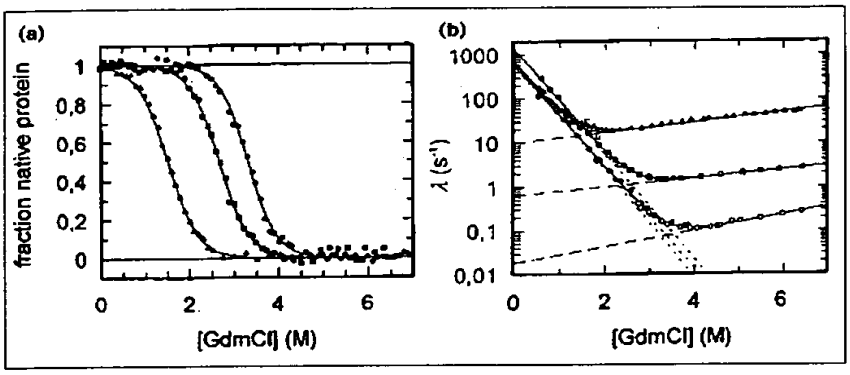
Hypothetical temperature profile of the free energy of **(a)** mesophilic and **(b–d)** thermophilic proteins. ΔG is defined as the difference in the free energies between the native and denatured proteins. T_m and T'_m are the melting temperatures of the mesophilic and thermophilic variants, respectively. The minimum of the ΔG parabola for a given protein (i.e. maximum stability) is observed at a temperature that is much below the optimal growth temperature (T_{opt} and T'_{opt}) of the respective mesophilic or thermophilic organism.

heat and cold denaturation (Figure 1). Commonly, the latter becomes detectable only under moderately destabilizing conditions [14,15•]. In the case of proteins from thermophiles, the ΔG versus temperature profile is either flattened or increased to larger $\Delta G_{N\rightarrow U}$ levels, rather than being shifted to higher temperatures. The ΔG maximum is always far below the optimal growth temperature; this holds also true for the (hyper-)thermophilic proteins [12•,16].

In order to simulate the effect of temperature on folding, the *in vitro* denaturation/renaturation of hyperthermophilic glyceraldehyde 3-phosphate dehydrogenase (GAPDH) was studied at 0–100°C. Refolding over a wide temperature range was found to yield the native state, even beyond the physiological temperature range, indicating that thermal stability refers not only to the native state, but also to intermediates on the folding pathway, independent of their states of association. At 0°C, the hyperthermophilic enzyme is trapped as a tetrameric intermediate with molten globule-like properties; upon shifting the temperature beyond ~10°C, the native state is reached instantaneously [16,17].

Regarding the folding kinetics, available data allow the conclusion that increasing intrinsic stability is reflected by a decrease in the rate of unfolding. In this context, mutant studies have shown that enhanced stability may be determined kinetically rather than thermodynamically [18•]. Comparing the unfolding and folding kinetics of the all-β, single-domain cold-shock proteins from *Bacillus subtilis*,

Figure 2



The conservation of the unfolding/folding mechanism of cold-shock proteins. **(a)** Equilibrium unfolding transitions of cold-shock proteins from *B. subtilis* (▲), *B. caldolyticus* (■) and *T. maritima* (●) induced by guanidinium chloride (GdmCl) at 25°C and monitored by intrinsic fluorescence. Least-squares fit analyses based on the two-state model U ⇌ N (full lines) yield stabilization energies ΔG_{stab} = 11.3, 20.1 and 26.2 kJ/mol for *B. subtilis*, *B. caldolyticus* and *T. maritima* Csp, respectively. **(b)** Kinetics of unfolding (open symbols) and refolding (closed symbols) of *B. subtilis* (△,▲), *B. caldolyticus* (□,■) and *T. maritima* Csp (○,●), respectively. The apparent rate constant, λ, is plotted against the concentration of GdmCl. The fits are on the basis of the linear two-state model. Reproduced with permission from [20•].

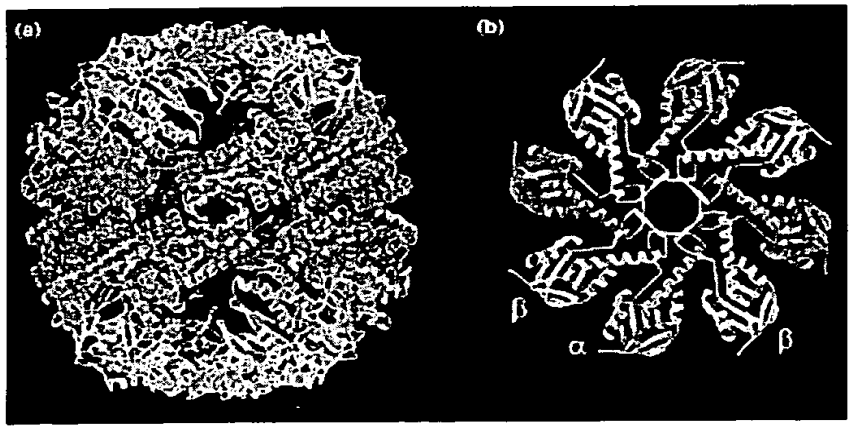
B. caldolyticus and *Thermotoga maritima* (Tm) (denaturation temperature T_m = 52, 72 and ~90°C, respectively), unfolding was shown to exhibit significantly slower kinetics with strongly decreasing rates, in accordance with the increase in stability. On the other hand, folding occurred extremely fast (τ ~1 ms), in a simple two-state reaction with closely similar kinetics, despite numerous sequence variations among the three proteins (Figure 2) [19,20•].

In the case of large proteins, generally, kinetic partitioning, that is, aggregation as a side reaction, competes with proper folding [11]. As a consequence, due to (partial) irreversibility, neither chemical denaturants nor temperature allow thermodynamic stability data to be determined. This is especially true for multidomain proteins and, to an even greater extent, for oligomers and multifunctional fusion proteins [16,18•,21,22•]. In certain cases, full reversibility was accomplished by choosing appropriate solvent conditions [23]. Evidently, kinetic partitioning *in vivo* is taken care of by molecular chaperones, if it is allowed to interpret the heat-stress response of, for example, *Pyrodictium occultum*, *Thermus thermophilus*, *Sulfolobus shibatae* and *Methanopyrus kandleri* along this line. All four species have been shown to express ATP-dependent 'thermosomes', sometimes to extreme levels [24,25]. The three-dimensional structure of these thermosomes exhibits a relationship with bacterial and eukaryal chaperones, with the only distinction being that the hyperthermophilic archaeal thermosomes obviously do not require co-chaperones; in the case of the thermosome from *T. acidophilum*, the apical domain seems to adopt this function [26•,27••] (Figure 3).

Forces and mechanisms involved in protein stabilization

The spatial structure of proteins is determined by electrostatic forces between polar and ionized groups and by hydrophobic effects involving nonpolar residues. The

Figure 3



Structure of the hexadecameric *T. acidophilum* thermosome (class II) seen in **(a)** side view and **(b)** top view, illustrating the 'cage-like' assembly with its eightfold symmetry. The apical domains close off both end cavities. Protrusions extend centrally from each thermosome apical domain, forming side-by-side contacts and a central β-sheet ring, comprising a built-in equivalent of GroES in the GroE system. Domains are colored in red (equatorial), green (intermediate) and yellow (apical). Within each complex, domains of aligned subunits are highlighted in blue (equatorial), light blue (intermediate) and violet (apical). (b) Thermosome α (red/violet) and β (yellow) apical domains. Reproduced with permission from [38••].

physical nature of the latter was recently interpreted as being entropic and enthalpic due to significant contributions from van der Waals' forces [1,8,28]. Attempts to obtain quantitative estimates of the different types of weak interactions have been based both on a comparison of known X-ray structures of mesophilic and extremophilic homologs, and on thermodynamic studies of point mutants. The $\Delta\Delta G_{N\rightarrow U}$ for extremophilic proteins is of the same order of magnitude as the overall free energy of stabilization $\Delta G_{N\rightarrow U}$ observed for mesophilic proteins. The following conclusions may be drawn regarding the significance of the various types of interactions:

1. $\Delta G_{N\rightarrow U}$ is equivalent to the energy required to break, at most, five hydrogen bonds, corresponding to about 1% of their total number in the folded structure of an average single-domain protein [11,29].

2. In the unfolded state, a 10 kDa protein exposes about 440 polar sites, half of them involved in internal hydrogen bonds in the native state. As a consequence, even a marginal difference in hydrogen-bond strength between water–water and water–protein hydrogen bonds will be magnified to an energy change that may well exceed $\Delta G_{N\rightarrow U}$ [29].

3. Water release from polar and nonpolar sites leads to an increase in entropy, which is supposed to be the driving force in many folding and assembly processes [30].

4. α helices and extended β structures contribute significantly to protein stability; in this context, helix-dipole interactions with charged groups in their vicinity are highly significant [9,31].

5. Contributions may also come from multiple hydrogen bonds between arginine groups and backbone carbonyl oxygens [32].

6. Only about 70% of the theoretically available hydrophobic contributions are realized as a consequence of the balance of favorable contributions to $\Delta G_{N\rightarrow U}$ on protein folding, leaving ample space for additional optimization [29].

7. As charged groups are commonly exposed to the aqueous solvent, intramolecular coulombic interactions cannot be of major importance in protein stabilization, unless they form clusters. Most of the polar sites in the cores of proteins are internally hydrogen bonded. The lack of this kind of internal saturation is found to be strongly destabilizing unless there are structured water molecules available to compensate for geometrical constraints [9,33]. Such 'ordered clathrate hydrates' have also been reported for patches of nonpolar residues exposed to the aqueous medium. Whether and how they contribute to stability is still unresolved [34].

The effects of temperature, pressure, charge and water activity on weak intermolecular interactions are highly complex. In the case of temperature, direct measurements of intermolecular forces have shown that the energy per polar group can exceed the thermal energy [35]. Hydrogen bonds are favored at low temperature and become weaker as the temperature is increased. Due to the compensatory effects in the total energy balance, predictions with respect to the significance of any specific type of interaction cannot be made.

Regarding the structural levels contributing to protein stability, increments may originate from local nearest neighbor interactions, secondary and supersecondary structural elements, subdomains, domains and subunits. The cumulative effect in terms of, at least, qualitative additivity has been illustrated by fragment and mutant studies [9,10,21,36,37]. At the quaternary level, a remarkable feature of hyperthermophiles is the occurrence of anomalous states of association and fused multifunctional proteins. Obviously, the reduction in the accessible surface area is associated with extreme thermophilicity [38••,39–47,48••,49]. Phosphoribosyl anthranilate isomerase (TmPRAI) and the bifunctional phosphoglycerate kinase–triosephosphate isomerase (TmPGK–TIM) fusion protein from *T. maritima* may serve as examples [39,47]. TmPRAI is a dimer with a complete $(\alpha\beta)_8$-barrel fold. In the monomeric *Escherichia coli* enzyme, the α_5 helix is replaced by a loop. The increase in stability comes from two long protrusions, which fit into cavities in the other subunit, thus favoring dimerization. Moreover, the sidechains of the N-terminal methione and the C-terminal leucine are immobilized in a hydrophobic cluster and, finally, the number of ion pairs is increased [48••]. In the case of the TmPGK–TIM fusion protein, a (–1) frameshift leads to the expression of both monomeric 43 kDa TmPGK and tetrameric 286 kDa TmPGK–TIM [39]. The dissection of the gene and the subsequent cloning, expression and characterization of the separate TmTIM showed that, in holding the complex together through three-dimensional domain swapping [50], the dimer of dimers gains significant stability within the fusion protein [12•,21,22•,51•]. The isolated entities are stabilized by the above increments. From a structural point of view, TmPGK in the fusion protein is practically indistinguishable from its thermophilic and mesophilic counterparts [39,51•,52]. In contrast, TmTIM (as the core of the fusion protein) differs significantly as a result of the additional intersubunit contacts (D Maes, RK Wierenga, personal communication). As one would expect for a complex multidomain oligomer, the folding of the fusion protein is hampered by kinetic partitioning. *In vivo*, folding might be assisted by thermosomal chaperone activity. Lactate dehydrogenase (LDH) has been used as a paradigm to summarize the various strategies of intrinsic stabilization. In this case, octamer formation is an 'artifact' of overexpression in *E. coli*, rather than a means of stabilization [53••,54]. Apart from intrinsic protein stabilization and

Table 1

The relative amino acid compositions of mesophiles and thermophiles [38••].

Amino acid*	Mesophiles	Thermophiles
Charged residues (DEKRH)	24.11%	29.84%
Polar/uncharged residues (GSTNQYC)	31.15%	26.79%
Hydrophobic residues (LMIVWPAF)	44.74%	43.36%

*One-letter abbreviations of the amino acid residues are given in brackets.

chaperoning, extrinsic factors, such as ions, carbohydrates or cofactors, may have additional stabilizing effects [7,55].

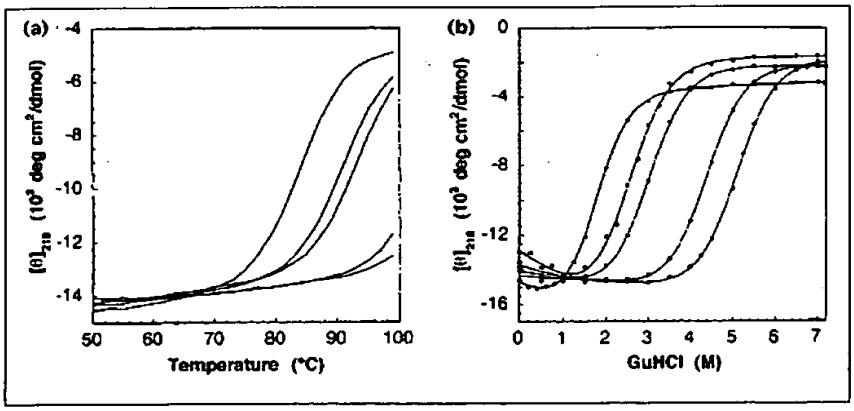
Genomes, phylogeny and general 'rules' of extremophilic adaptation

Inspecting available genome sequences and protein structures, there seems to be no way to unambiguously correlate extremes of physical conditions with either the amino acid composition or the three-dimensional structure of a given set of homologous proteins. What has been gained with increasing detail is an understanding of the phylogeny of microorganisms and their genes [56].

16S ribosomal RNA has been used for more than a decade to analyze evolutionary relationships between organisms [57••]. As a result, the three domains of life — bacteria, archaea and eukarya — were defined, with the general conclusion being that bacterial and archaeal hyperthermophiles are close to the root of the phylogenetic tree, preceding their mesophilic counterparts. At this point, it is important to note that such phylogenetic relationships say nothing about the temperature at which life started [58]. The advent of complete genome sequences made it clear that the phylogenetic tree has more complex roots than expected so far. Considering one single genome, single genes may or may not agree with the rRNA tree. Even more perplexing is the fact that genomes contain a mix of DNAs, some are close to archaea, while others are close to bacteria. The recently determined complete genome sequence of the hyperthermophilic bacterium *Aquifex aeolicus* [38••] shows the phenomenon in a unique way. A comparison of several of its genes with their counterparts in a range of species from all three domains of life reveals that *Aquifex* gets different pylogenetic placements, depending on which gene is being considered. It looks as if each gene has its own history, possibly due to lateral gene transfer or the 'swapping of genes' among organisms. Although the mechanism of gene swapping is still unknown, there seems to be no better explanation for the observation that 17 out of 34 families of eukaryotic proteins that date back to early cell evolution look as if they come from bacteria, while only eight families show a greater similarity to archaea, the supposed ancestor of eukarya. In spite of these inconsistencies, presently available genome sequences still fit the three-domain hypothesis [59•].

Among the 14 bacterial and archaeal genomes sequenced so far, five belong to hyperthermophiles: *Aquifex aeolicus*, *Archaeoglobus fulgidus*, *Methanobacterium thermoautotrophicum*, *Methanococcus jannaschii* and *Pyrococcus horikoshii* (http://www.ncbi.nlm.nih.gov/Entrez/Genome/org.html). A comparison of these genomes with respect to specific genes from mesophiles provides a data set that is sufficiently large to extract certain trends in amino acid usage (Table 1). Some correlations seem to hold — compared to mesophiles, genomes of thermophiles encode higher levels of charged amino acids, primarily at the expense of uncharged polar residues. Glutamine, in particular, seems to be significantly discriminated against in hyperthermophiles (compare [60]). This observation might be rationalized by an increased rate of deamidation of this residue at higher temperatures. One might expect the same difference in the number of asparagines; however, this residue does not appear to be subject to similar discrimination [38••].

Figure 4



Circular dichroism measurements of the *Streptococcal* protein Gβ1 domain and its mutants with optimized core packing, a decreased hydrophobic surface area and improved secondary structure propensity. For the nomenclature of the mutants, 'c' refers to residues in the core, whereas 'b' stands for bordering residues. Profiles from left to right refer to: Gβ1 (wildtype); Gβ1-c3 (Y3F, L7I, V39I); Gβ1-c3b1 (Y3F, L7I, V39I, T25E); Gβ1-c3b2 (Y3F, L7I, V39I, T61I, T18I); Gβ1-c3b4 (Y3F, L7I, V39I, T61I, T18I, T25E, V29I). **(a)** Thermal unfolding. **(b)** Chemical denaturation by guanidinium chloride at 50°C [66••]. Reproduced with permission from [66••].

Earlier attempts to find 'traffic rules' of protein stabilization were based on sequence comparisons in relatively small sample sizes. A critical analysis proved these early predictions to be statistically insignificant [61]. Examining the increasing number of high-resolution structures, a variety of strategies of thermal adaptation have emerged [62–64]. Recently, Vogt and Argos [65], in an attempt to rank the strategies proposed so far, examined 16 protein families, each with at least one known thermophilic and one known mesophilic structure, focusing on hydrogen bonds, ion pairs, polar surface composition, internal cavities, packing densities and secondary structure composition. The results show that enhanced thermostability is correlated with a consistent increase in the number of hydrogen bonds and ion pairs, apart from an increase in polar surface area.

The picture that is now emerging of the stabilization effects of hyperthermophilic proteins, showing a mosaic of different strategies of thermal adaptation, has been tested and supported by the prediction of mutagenesis effects, as, for example, in the case of the *Streptococcal* protein Gβ1 [66••]. A sevenfold mutant of the protein was designed by an objective computer algorithm that modeled in parallel several complex contributions to stability into the protein, incorporating optimized core packing, increased burial of hydrophobic surface area, more favorable helix-dipole interactions and the improvement of secondary-structure propensities. The stabilization effects were found to be additive, with a ΔΔG of 18 kJ/mol and a shift in the thermal transition from 83°C for the wildtype protein to more than 99°C for the sevenfold mutant (Figure 4).

In a limited number of cases, enzymes from psychrophiles have been used in order to expand the temperature scale and confirm present ideas on protein stabilization. Summarizing the state-of-the-art, it appears that the adaptive mechanisms involve weakening of intramolecular interactions and/or increasing interactions with the solvent; both these tendencies lead to enhanced flexibility, that is, catalysis at a lower energy cost [67•]. TIM [68•], 3-isopropylmalate dehydrogenase (IPMDH) [69] and citrate synthase (CS) [70] have been studied in attempts to recognize specific adaptive strategies. In the latter two cases, a comparison of the amino acid compositions, as well as sequence alignments and homology modeling, have shown that there are no traffic rules, except for an increase in extended surface loops and a decrease in the number of proline residues in loops, equivalent to an increase in the configurational entropy of the denatured state. The observed changes in charge distribution are difficult to classify.

A pairwise comparison of homologous proteins with different thermal stabilities has been widely applied in order to discover strategies of thermal adaptation. Based on known three-dimensional structures, stabilizing features were deduced. In order to test the proposals, site-directed

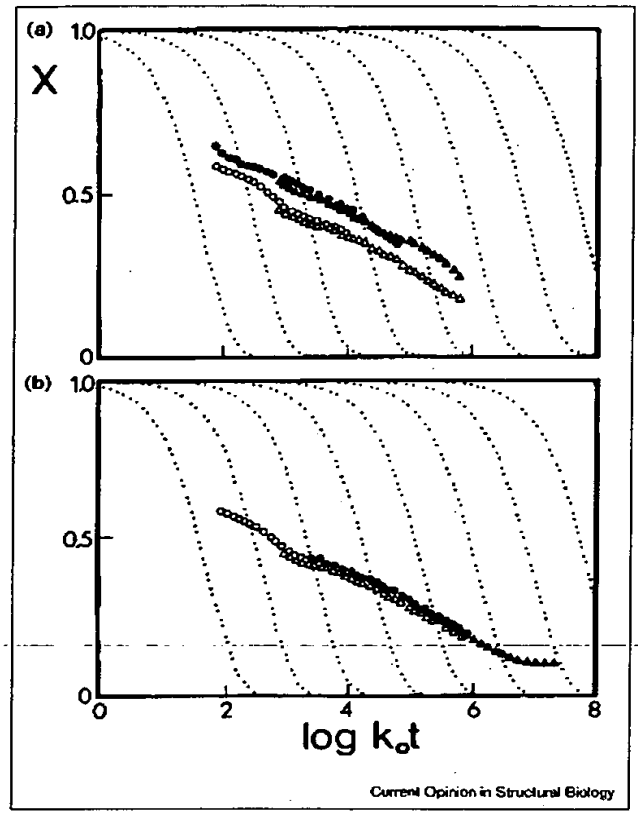
Table 2

Strategies of thermal stabilization of selected thermophilic enzymes.

Protein	Major cause(s) of thermostability*
CS (*P. furiosus*)	Increased compactness, enhanced subunit interactions, increased number of intersubunit ion pairs, shortening of loops [88••].
Ferredoxin (*T. maritima*)	Structurally: stabilization of α helices; replacement of conformationally strained residues by glycines; strong docking of the N-terminal methionine; increase in the number of hydrogen bonds [89]. Thermodynamically: flat ΔG versus temperature profile caused by a low ΔCp of unfolding [90•].
Ferredoxin (*Synechococcus elongatus*)	Extension of the hydrophobic core, a unique hydrophobic patch on the surface of a β sheet, two unique ion-pair networks [91].
GluDH (*P. furiosus*)	A series of extended ion-pair networks on protein subunit surfaces and ion-pair networks buried at interdomain and intersubunit interfaces, enhanced packing within the inner core, amino acid replacements increasing the hydrophobicity and sidechain branching (Val→Ile) [92,93•].
GAPDH (*T. maritima*)	Large number of additional salt bridges [94].
Indole 3-glycerol phosphate synthase (*Sulfolobus solfataricus*)	Large number of additional salt bridges (partly cross-linking adjacent helices), increased helix capping, dipole stabilization, increased hydrophobic interactions, strengthening of chain termini and solvent-exposed loops [95].
IPMDH (*T. thermophilus*)	Increased number of ion pairs and hydrogen bonds, extended hydrophobic subunit interactions and improved packing of the hydrophobic core, shortened chain termini [96••,97•,98•].
LDH (*T. maritima*)	Increased number of ion pairs, decreased hydrophobic surface area, increased helicity, less cavity volume [53••].
PGK (*T. maritima*)	Increased rigidity by additional ion pairs, stabilization of α helix and loop regions, stabilization by fusion with TIM [23,51•,52].

*Listed according to the authors' priorities (if described). ΔCp, heat capacity change.

Figure 5



Hydrogen–deuterium exchange of GAPDH from *T. maritima* (closed symbols) and rabbit (open symbols), measured at pH 6.0 (●,○) and 7.0 (▲,△) plotted as relaxation spectra. For experimental details, see [72]. **(a)** Measurements at constant temperature (25°C); increasing X values reflect increased rigidity. **(b)** Measurements at 68°C for TmGAPDH and 25°C for rabbit GAPDH. Coincidence of the curves indicates similar flexibility (P Závodszky, C Böthe, J Kardos, A Svingor, R Jaenicke, unpublished data).

mutagenesis or similar control experiments were performed. Examples of strategies of thermal adaptation are given in Table 2.

On going from lower to higher growth temperatures, numerous differences are most frequently reported: the clustering of (intrasubunit and/or intersubunit) ion pairs; improved packing of the hydrophobic core (increased van der Waals' interactions); additional networks of hydrogen bonds and enhanced secondary structure propensity; increased helix-dipole stabilization; an increased polar surface area; a decreased number and total volume of cavities; and burying hydrophobic surface area by either tightening interdomain and intersubunit contacts or by increasing the state of association. For a collection of 'recipes' to improve thermal stability, see [71]. The common denominator of all the adaptive changes is the conservation and optimization of the functional state of the given protein. Experimentally, this can be defined either by biological activity or by physical parameters connected with the flexibility or rigidity of the polypeptide chain. A suitable approach is hydrogen–deuterium exchange kinetics at various temperatures. They confirm the hypothesis that under optimum physiological conditions, proteins are in corresponding states [1,72,73] (Figure 5).

Halophiles

Halophilic proteins require multimolar salt concentrations for activity and denature at low salt. Malate dehydrogenase from *Haloarcula marismortui* was the first extremely halophilic enzyme structure to be solved by X-ray crystallography, ending a long physicochemical odyssey from a 'dimeric core with loops' to the real tetrameric 'LDH homolog' [74]. Anomalous features of the protein include an excess of acidic residues over basic residues and an increase in the number of intramolecular salt bridges compared to nonhalophilic homologs [75]. An explanation for its halophilicity is its tendency to increase water binding to the protein surface, in order to compete with the highly concentrated salt solution for water of hydration. This idea was confirmed by the significant increase in peripheral hydrogen bonds in the case of the 1.9 Å X-ray structure of 2Fe–2S ferredoxin from *H. marismortui* [76]. On the other hand, the observed significant increase in halophilicity after mutating a single glutamic acid position (unique in halophilic malate dehydrogenases) to arginine sheds some doubt on the hypothesis [77]. An additional feature of halophilic adaptation, which emerged from homology modeling of glutamate dehydrogenase (GluDH) from *Halobacterium salinarium* (Hs), is a significant reduction in 'exposed hydrophobic character', due to a decrease in the number of surface-exposed lysine residues [78•]; whether this finding will stand the experimental test of X-ray analysis remains to be seen.

Considering the X-ray structure of a moderately halophilic protein, monomeric dihydrofolate reductase (DHFR) from *Haloferax volcanii*, no striking differences from its mesophilic counterpart have been detected, except for some charge clusters of negatively charged residues [79•]. Similar observations were reported for the dimeric dihydrolipoamine dehydrogenase from the same source. In this case, site-directed mutagenesis of four coordinated glutamate residues involved in a charge cluster at the subunit interface was shown to have a profound effect on the salt dependence of the activity of the enzyme [80•].

In summarizing available data, halophilic enzymes do not exhibit specific structural properties. The most notable feature is the presence of clusters of negatively charged residues; their repulsion at low salt may account for the instability of halobacterial proteins at low salt. A similar structural motif, with the associated salt effect on stability, is also present in nonhalophilic proteins, such as ribonuclease T1 [81].

Conclusions

The significant enhancement in our understanding of extremophilic adaptation has lately come from four directions: the discovery of more extreme extremophiles; complete genome sequences; new genetic methods; and rapidly increasing numbers of high-resolution, three-dimensional structures of proteins. Regarding thermophiles, a maximum growth temperature of 113°C has been the latest news for the Guinness Book of Records [82]; apart from this minor sensation, successful expeditions to the abyssal depth of Challenger Deep brought thousands of microorganisms to the surface, although very few of them were barophilic [2,3•]. Time will show whether any common properties exist within the broad barotolerant microbial population and its nucleic acids and proteins. The second direction has opened new horizons into phylogenetics, both unveiling incongruities of the phylogenetic tree, from its root to its major branchings, and generating new scenarios that will soon lead to a boom of experiments both in genetics and in biocomputing. The third point, that is, new techniques, such as directed evolution, will become essential, because simulating evolution needs fast selection methods, which are becoming increasingly available. Finally, X-ray analysis has become a common tool in many laboratories, so the number of high-resolution structures is growing at an unexpected pace. It may sooner or later reach the level at which the sample size becomes sufficiently large for significant statistical analyses. With the increase in information from the regime of thermophiles, a limited number of strategies of stabilization has emerged, which may be used for careful predictions, as well as in industrial applications. A recent illustration is the cumulative mutation approach underlying the engineering of a thermolysine-like protease that combines low temperature activity with thermal stability beyond the boiling point of water [83]. In this context, the potential of chemistry both to expand the protein alphabet by including noncanonical amino acids or to increase chelating affinity for extrinsically stabilizing ligands promises to be a gold mine for future protein chemistry and bioengineering studies [84–87].

Restricting ourselves to the canonical amino acids, thermostability has been shown to be the cumulative effect of packing efficiency (mainly through van der Waals' interactions), networks of ion pairs and/or hydrogen bonds (including α-helix stabilization), the reduction of conformational strain (loop stabilization) and resistance to chemical modification. Comparing mesophilic and thermophilic enzymes, the increase in conformational rigidity suggests that evolutionary adaptation tends to maintain corresponding states with respect to conformational flexibility, that way optimizing biological function under specific physiological conditions. In halophiles, the correspondence is assumed to refer to the state of hydration of the protein. The complexity of multicomponent thermodynamics does not, however, allow clear-cut adaptive strategies to be defined. The same holds true for barophiles, acidophiles and alkalophiles.

Obviously, extremophilic adaptation refers to proteins in their native state, as well as in their nascent state. Since misfolding is expected to be an important side reaction under extreme conditions, the discovery of molecular chaperones in hyperthermophiles did not come as a surprise. Given their complex assembly structure, however, the determination of their spatial structure at high resolution has been a most amazing success. The elucidation of their function as a cage or an excluded volume component might be the subject of a future article in *Current Opinion in Structural Biology* dealing, for example, with hyperthermophilic heat-shock proteins.

Acknowledgment

This review is dedicated to Professor Alfred Schellenberger on the occasion of his 70th birthday.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Jaenicke R: **Protein stability and molecular adaptation to extreme conditions.** *Eur J Biochem* 1991, **202**:715-728.

2. Takami H, Akira I, Fuji F, Horikoshi K: **Microbial flora in the deepest sea mud of the Mariana trench.** *FEMS Microbiol Lett* 1997, **152**:279-285.

3. • Kato C, Li L, Nogi Y, Nakamura Y, Tamaoka J, Horikoshi K: **Extremely barophilic bacteria isolated from the Mariana trench, Challenger Deep, at a depth of 11,000 meters.** *Appl Environ Microbiol* 1998, **64**:1510-1513.

Obligatory barophilic proteobacteria (genus *Moritella*) with optimal growth at more than 70 MPa and no growth at less than 50 MPa were isolated from a sample taken from the Mariana Trench at a depth of 10,898 m.

4. Jaenicke R: **What ultrastable globular proteins teach us about protein stability.** *Biochemistry (Moscow)* 1998, **63**:312-321.

5. Bernhardt G, Lüdemann H-D, Jaenicke R, König H, Stetter KO: **Biomolecules are unstable under 'Black smoker' conditions.** *Naturwissenschaften* 1984, **71**:583-586.

6. Leibrock E, Bayer P, Lüdemann H-D: **Non-enzymatic hydrolysis of ATP at high temperatures and high pressures.** *Biophys Chem* 1995, **54**:175-180.

7. Carpenter JF, Clegg JS, Crowe JH, Somero GN (Eds): **Compatible solutes and macromolecular stability.** *Cryobiology* 1993, **30**:201-241.

8. Dill KA: **Dominant forces in protein folding.** *Biochemistry* 1990, **29**:7133-7155.

9. Matthews BW: **Structural and genetic analysis of the folding and function of T4 lysozyme.** *FASEB J* 1996, **10**:35-41.

10. Jaenicke R: **Stability and folding of domain proteins.** *Progr Biophys Mol Biol* 1999, in press.

11. Jaenicke R: **Protein folding and association: *in vitro* studies for self-organization and targeting in the cell.** *Curr Top Cell Reg* 1997, **34**:213-334.

12. • Grättinger M, Dankesreiter A, Schurig H, Jaenicke R: **Recombinant PGK from the hyperthermophilic bacterium *Thermotoga maritima*: catalytic, spectral and thermodynamic properties.** *J Mol Biol* 1998, **280**:525-533.

As derived from the temperature dependence of $\Delta G_{N\rightarrow U}$, optimum stability is observed at 30°C. The extrapolated temperatures of cold and heat denaturation are −10 and +85°C, respectively. Comparing the free energy of stabilization of the *T. maritima* enzyme with its counterpart from yeast, ΔΔG amounts to 85 kJ/mol.

13. Pfeil W: *Protein Stability and Folding: a Collection of Thermodynamic* • *Data.* Heidelberg: Springer Berlin; 1998.
This book makes the vast amount of data in the literature from 1964 to 1996 accessible to energy calculations of natural and modified proteins, tabulating Gibbs energy changes, enthalpy changes and heat capacity changes.

14. Franks F: **Protein destabilization at low temperatures.** *Adv Protein Chem* 1995, **46**:105-139.

15. Graziano G, Catanzano F, Riccio A, Barone G: **A reassessment of** • **the molecular origin of cold denaturation.** *J Biochem* 1997, **122**:395-401.
Based on Ikegamis' statistical thermodynamic model of the structural fluctuations of proteins, Gibbs energy changes associated with the transfer of *N*-alkyl-amides to water were used to devise a model of protein stability. The results indicate that the amphiphilic nature of the polypeptide chain and the presence of water are fundamental to cold denaturation.

16. Jaenicke R: **GAPDH from *Thermotoga maritima*: strategies of protein stabilization.** *FEMS Microbiol Rev* 1996, **18**:215-224.

17. Jaenicke R, Schurig H, Beaucamp N, Ostendorp R: **Structure and stability of hyperstable proteins: glycolytic enzymes from hyperthermophilic bacterium *Thermotoga maritima*.** *Adv Protein Chem* 1996, **48**:181-269.

18. Pappenberger G, Schurig H, Jaenicke R: **Disruption of an ionic** • **network leads to accelerated thermal denaturation of GAPDH from the hyperthermophilic bacterium *Thermotoga maritima*.** *J Mol Biol* 1997, **274**:676-683.
Replacing the central arginine residue of an ionic network between the N and C-terminal helices (αB and α3) by alanine or asparagine causes a shift of 7°C in the temperature of half-denaturation. The decrease in the free energy of activation $\Delta G^{\neq}$ for thermal unfolding amounts to 4 kJ/mol (at 100°C); at physiological temperature, this yields a strong increase of the half-life of the enzyme.

19. Welker C, Böhm G, Schurig H, Jaenicke R: **Cloning, overexpression, purification and physicochemical characterization of a cold shock protein homolog from *Thermotoga maritima*.** *Protein Sci* 1998, in press.

20. Perl D, Welker C, Schindler T, Schröder K, Marahiel MA, Jaenicke R, • Schmid FX: **Conservation of rapid two-state folding in mesophilic, thermophilic, and hyperthermophilic cold shock proteins.** *Nat Struct Biol* 1998, **5**:229-235.
Cold-shock proteins from *B. subtilis*, *B. caldolyticus* and *T. maritima* differ significantly in both their sequences and thermal stabilities ($\Delta G_{N\rightarrow U} = 11$, 20 and 26 kJ/mol, respectively), but show similar, extremely fast folding kinetics ($k_{U\rightarrow N} = 690$, 1370, 565 s^{-1}). The lack of correlation between the thermodynamic stability and folding rate indicates that proteins with a more stable folded state do not necessarily fold faster, as suggested from lattice models.

21. Beaucamp N, Schurig H, Jaenicke R: **The PKG-TIM fusion protein from *Thermotoga maritima* and its constituent parts are intrinsically stable and fold independently.** *Biol Chem* 1997, **378**:679-685.

22. Beaucamp N, Hofmann A, Kellerer B, Jaenicke R: **Dissection of the** • **gene of the bifunctional PGK-TIM fusion protein from the hyperthermophilic bacterium *Thermotoga maritima*: design and characterization of the separate TIM.** *Protein Sci* 1997, **6**:2159-2165.
The structural properties of recombinant separate *T. maritima* triosephosphate isomerase (TmTIM) resemble those of its mesophilic homologs. Its catalytic activity at 80°C comes close to that of higher vertebrates at their physiological temperature, in accordance with the idea that, under *in vivo* conditions, enzymes occupy 'corresponding states'. Within the tetrameric phosphoglycerate kinase (PGK)–TIM fusion protein, the intrinsic thermal stability of TmTIM is enhanced ($\Delta T_m = 8$°C); the same is true for the catalytic efficiency.

23. Zaiss K, Schurig H, Jaenicke R: **Differential scanning calorimetry of PGK from the hyperthermophilic bacterium *Thermotoga maritima*.** In *Biocalorimetry: Application of Calorimetry in Biological Sciences.* Edited by Ladbury J, Chowdhry BZ. Sussex: John Wiley; 1998:283-293.

24. Trent JD, Kagawa HK, Yaoi T, Olle E, Zaluzec NJ: **Chaperonin filaments: the archaeal cytoskeleton?** *Proc Natl Acad Sci USA* 1997, **94**:5383-5388.

25. Minuth T, Frey G, Lindner P, Jaenicke R, Stetter KO: **Recombinant homo- and hetero-oligomers of an ultrastable chaperonin from *Pyrodictium occultum* show chaperone activity *in vitro*.** *Eur J Biochem* 1998, in press.

26. Klumpp M, Baumeister W, Essen L-O: **Structure of the substrate** • **binding domain of the thermosome, an archaeal group II chaperonin.** *Cell* 1997, **91**:263-270.
Based on the 2.3 Å crystal structure of the (Ser214–Asn365) substrate-binding domain at the α subunit from the *T. acidophilum* thermosome and molecular modeling of the hexadecameric complex from the tetradecameric GroEL structure, a structural model of the intact thermosome was presented; the model is consistent with cryoelectron microscopy data.

27. Ditzel L, Löwe J, Stock D, Stetter KO, Huber H, Huber R, Steinbacher •• S: **Crystal structure of the thermosome, the archaeal chaperonin and homolog of CCT.** *Cell* 1998, **93**:125-138.
The determination of the structure of the complete hexadecameric thermosome from *T. acidophilum* at 2.6 Å resolution confirms its topological homology with the GroEL system, with a novel type of inter-ring contact. Parts of the apical domain form a lid, substituting for the absent GroES-like chaperonin 10. The central cavity has a polar surface that is involved in folding assistance. The active site geometry allows a mechanism for ATP hydrolysis to be proposed.

28. Makhatadse GI, Privalov PL: **Energetics of protein structure.** *Adv Protein Chem* 1996, **47**:307-425.

29. Pace CN, Shirley BA, McNutt M, Gajiwala K: **Forces contributing to the conformational stability of proteins.** *FASEB J* 1996, **10**:75-83.

30. Jaenicke R: **Folding and association of proteins.** *Progr Biophys Mol Biol* 1987, **49**:117-237.

31. Colacino F, Crichton RR: **Enzyme thermostabilization: the state of the art.** *Biotechnol Genet Eng Rev* 1997, **14**:211-277.

32. Borders CL Jr, Broadwater JA, Bekeny PA, Salmon JE, Lee AS, Elridge AM, Pett VB: **A structural role for arginine in proteins: multiple H-bonds to backbone carbonyl oxygens.** *Protein Sci* 1994, **3**:541-548.

33. Stigter D, Dill KA: **Charge effects on folded and unfolded proteins.** *Biochemistry* 1990, **29**:1262-1271.

34. Teeter MM: **The water structure surrounding proteins.** In *Protein Folding.* Edited by Gierasch L, King J. Washington DC: AAA; 1990:44-54.

35. Leikin S, Parsegian VA: **Temperature-induced complementarity as a mechanism for biomolecular assembly.** *Proteins* 1994, **19**:73-76.

36. Matthews BW: **Studies on protein stability with T4 lysozyme.** *Adv Protein Chem* 1995, **46**:249-278.

37. Vita C, Fontana A, Jaenicke R: **Folding of thermolysin fragments: hydrodynamic properties of isolated domains and subdomains.** *Eur J Biochem* 1989, **183**:513-518.

38. Deckert G, Warren PV, Gaasterland T, Young WG, Lenox AL, Graham •• DE, Overbeek R, Snead MA, Keller M, Aujay M *et al.*: **The complete genome of the hyperthermophilic bacterium *Aquifex aeolicus*.** *Nature* 1998, **392**:353-358.
According to 16S rRNA analysis, *A. aeolius* is one of the earliest diverging bacteria known. With 1,551,335 base pairs, its genome is only one-third the size of the *E. coli* genome. It contains 93% protein coding regions, so that most of the genes are expected to be related to known open reading frames characterized elsewhere. In spite of its hyperthermophilic nature, only a few specific indications of thermophilicity are apparent from the genome.

39. Schurig H, Beaucamp N, Ostendorp R, Jaenicke R, Adler E, Knowles J: **PGK and TIM from the hyperthermophilic bacterium *Thermotoga maritima* form a covalent bifunctional enzyme complex.** *EMBO J* 1995, **14**:442-451.

40. Gabelsberger J, Liebl W, Schleifer KH: **Purification and properties of recombinant β-glucosidase of the hyperthermophilic bacterium *Thermotoga maritima*.** *Appl Microbiol Biotechnol* 1993, **40**:46-52.

41. Kengen SW, Luesink EJ, Stams AJ, Zehnder AJ: **Purification and characterization of an extremely thermostable β-glucosidase from the hyperthermophilic archaeon *Pyrococcus furiosus*.** *Eur J Biochem* 1993, **213**:305-312.

42. Schurig H, Rutkat K, Rachel R, Jaenicke R: **Octameric enolase from the hyperthermophilic bacterium *Thermotoga maritima*: purification, characterization and image processing.** *Protein Sci* 1995, **4**:228-236.

43. Hess D, Krüger K, Knappik A, Halm P, Hensel R: **Dimeric PGKs from hyperthermophilic archaea. Cloning, sequencing and expression of the PGK gene of *Pyrococcus woesei* in *E. coli* and characterization of the protein. Structural and functional comparison with the PGK of *Methanothermus fervidus*.** *Eur J Biochem* 1996, **233**:227-237.

44. Kohlhoff M, Dahm A, Hensel R: **Tetrameric TIM from hyperthermophilic archaea.** *FEBS Lett* 1996, **383**:245-250.

45. Dams T, Böhm G, Auerbach G, Bader G, Schurig H, Jaenicke R: **Homo-dimeric dihydrofolate reductase from *Thermotoga maritima* shows extreme intrinsic stability.** *Biol Chem* 1998, **379**:367-371.

46. Aguilar CF, Sanderson I, Moracci M, Ciaramella M, Nucci R, Rossi M, Pearl LH: **Crystal structure of the β-galactosidase from the hyperthermophilic archaeon *Sulfolobus solfataricus*: resilience as a key factor in thermostability.** *J Mol Biol* 1997, 271:789-802.

47. Sterner R, Kleemann GR, Szadkowski H, Lustig A, Hennig M, Kirschner K: **PRAI from *Thermotoga maritima* is an extremely stable and active homo-dimer.** *Protein Sci* 1996, 5:2000-2008.

48. •• Hennig M, Sterner R, Kirschner K, Jansonius JN: **Crystal structure at 2.0 Å resolution of PRAI from the hyperthermophile *Thermotoga maritima*: possible determinants of protein stability.** *Biochemistry* 1997, 36:6009-6016.
In contrast to most mesophilic PRAIs, hyperthermophilic TmPRAI is a homodimer (as a *caveat*, in *E. coli*, the standard mesophile, PRAI is part of a bifunctional enzyme). In spite of the increased overall rigidity of TmPRAI, its activity at 80°C is about 35-fold higher compared to its mesophilic counterpart at 37°C, this way successfully competing with the spontaneous hydrolysis of the substrate at physiological temperature. Stability increments come from multiple hydrophobic interactions and additional ion pairs. Their elimination by protein engineering leads to the fully active monomer exhibiting drastically reduced kinetic stability (R Thoma, K Kirschner, personal communication).

49. Villeret V, Clantin B, Tricot C, Legrain C, Roovers M, Stalon V, Glansdorff N, van Beeumen J: **The crystal structure of *Pyrococcus furiosus* ornithine carbamoyltransferase reveals a key role for oligomerization in enzyme stability at extremely high temperatures.** *Proc Natl Acad Sci USA* 1998, 95:2801-2806.

50. Schlunegger MP, Bennett MJ, Eisenberg D: **Three-dimensional domain swapping.** *Adv Protein Chem* 1997, 50:61-122.

51. • Auerbach G, Huber R, Grättinger M, Zaiss K, Schurig H, Jaenicke R, Jacob U: **Closed structure of PGK from *Thermotoga maritima* reveals the catalytic mechanism and determinants of thermal stability.** *Structure* 1997, 5:1475-1483.
This 2.0 Å crystal structure determination adds to a series of four known phosphoglycerate-kinase-structures-from-mesophiles-and-moderate-thermophiles, allowing the structural determinants of thermostability to be analyzed. In addition, the abortive ternary complex (with the substrate and AMP–adenyl-imino diphosphate) gives a detailed picture of the hinge bending 'phosphoryl gripper' mechanism. Thermophilic characteristics include increased rigidity as a consequence of enhanced intramolecular interactions involving ion pairs, α helices and loop regions.

52. Auerbach G, Jacob U, Grättinger M, Zaiss K, Schurig H, Jaenicke R: **PGK from *Thermotoga maritima*.** *Biol Chem* 1997, 378:327-329.

53. •• Auerbach G, Ostendorp R, Prade L, Korndörfer I, Dams T, Huber R, Jaenicke R: **LDH from the hyperthermophilic bacterium *Thermotoga maritima*: the crystal structure at 2.1 Å resolution reveals strategies for intrinsic stabilization.** *Structure* 1998, 6:769-781.
Five lactate dehydrogenases (LDH) from moderate thermophiles and mesophiles (prokaryotes and eukaryotes) were compared with the structure of the quaternary complex of TmLDH (with NADH, the allosteric activator fructose 1,6-bisphosphate and the substrate analog oxamate as ligands). The three-dimensional folds of all five enzymes are very similar, allowing a correlation between structural parameters and thermal stability to be established: 7 out of 18 arginine residues are engaged in ion pairs and four form hydrogen bonds to mainchain carbonyl groups; the number of ion pairs correlates with the growth temperature; hydrogen bonds do not make a significant contribution; there is a decrease in hydrophobic surface area and an increase in helicity (including an additional 'thermo helix'), but no change in the amount of surface area buried in subunit contacts; TmLDH has fewer and smaller cavities than its less thermostable counterparts. The high cysteine content is not involved in metal binding; the relevance for stability of a four cysteine cluster is still to be explained.

54. Dams T, Ostendorp R, Ott M, Rutkat K, Jaenicke R: **Tetrameric and octameric LDH from the hyperthermophilic bacterium *Thermotoga maritima*: structure and stability of the two active forms.** *Eur J Biochem* 1996, 240:274-279.

55. Jaenicke R: **Oligomeric proteins.** In *Molecular Chaperones in the Life Cycle of Proteins*. Edited by Fink AL, Goto Y. New York: Marcel Dekker; 1998:35-70.

56. Koonin EV, Tatusov RL, Galperin ML: **Beyond complete genomes: from sequence to structure and function.** *Curr Opin Struct Biol* 1998, 8:355-363.

57. •• Woese C: **The universal ancestor.** *Proc Natl Acad Sci USA* 1998, 95:6854-6859.
In analogy to physical annealing, a 'genetic annealing model' of evolution is presented. In the beginning, simple cellular entities evolved with high rates of mutation and lateral gene transfer. Later, both rates dropped and finally approached the evolutionary dynamics existing today; organismic lineages still did not exist. The universal phylogenetic tree emerges as its peripheral branches develop. Thus, the universal ancestor is a diverse community of cells evolving as biological unit; over time, it was refined into a decreasing number of cell types, finally leading to the present 'kingdoms of life'.

58. Miller SL, Lazcano A: **The origin of life – did it occur at high temperatures?** *J Mol Evol* 1995, 41:689-692.

59. • Doolittle RF: **Microbial genomes opened up.** *Nature* 1998, 392:339-342.
Evidence from completely sequenced genomes seems to indicate that each gene has its own history due to the combined effects of mutations and lateral gene transfers. Available genome sequences (twelve bacteria and one yeast) are discussed in the context of phylogenetic trees, minimal gene content, the common ancestor and evolutionary mechanisms.

60. Catanzano F, Graziano G, Capasso S, Barone G: **Thermodynamic analysis of the effect of selective monodeamidation at Asp67 in ribonuclease A.** *Protein Sci* 1997, 6:1682-1693.

61. Böhm G, Jaenicke R: **On the relevance of sequence statistics for the properties of extremophilic proteins.** *Int J Peptide Prot Res* 1994, 43:97-106.

62. Russell RJM, Taylor GL: **Engineering thermostability: lessons from thermophilic proteins.** *Curr Opin Biotechnol* 1995, 6:370-374.

63. Querol E, Perez-Pons JA, Mozo-Villarias A: **Analysis of protein conformational characteristics related to thermostability.** *Protein Eng* 1996, 9:265-271.

64. Vieille C, Zeikus J: **Thermoenzymes: identifying molecular determinants of protein structural and functional stability.** *Trends Biotechnol* 1996, 14:183-191.

65. Vogt G, Argos P: **Protein thermal stability: hydrogen bonds or internal packing?** *Fold Des* 1997, 2:S40-S46.

66. •• Malakauskas SM, Mayo SL: **Design, structure and stability of a hyperthermophilic protein variant.** *Nat Struct Biol* 1998, 5:470-475.
The design algorithm models several proposed stability increments simultaneously, using empirical potential functions and a combinatorial optimization procedure based on the dead-end elimination theorem. The designed sevenfold mutant shows an increased stability (18 kJ/mol) and enhanced thermostability (> 99°C). This method may potentially be applied to the stabilization of other unrelated protein folds.

67. • Gerday C, Aittaleb M, Arpigny JL, Baise E, Chessa J-P, Garsoux G, Petrecu I, Feller G: **Psychrophilic enzymes: a thermodynamic challenge.** *Biochim Biophys Acta* 1997, 1342:119-131.
Although cold is the most widespread challenge in the biosphere, very little is known about cold adapted and psychrophilic proteins. The authors give an overview of 'cold enzymes' from Antarctic bacteria, focusing on α-amylase, β-lactamase, lipase and subtilisin.

68. • Alvarez M, Zeelen JP, Mainfroid V, Rentier-Delrue F, Martial JA, Wyns L, Wierenga RK, Maes D: **TIM of the psychrophilic bacterium *Vibrio marinus*.** *J Biol Chem* 1998, 273:2199-2206.
Comparing the crystal structures of *V. marinus* TIM (T_{opt} = 15°C) and *E. coli* TIM, the root mean square differences were found to be less than 0.6 Å, in accordance with the high sequence identity (66%). No conclusions can be drawn from the structural differences, except for residue Ala238 in *Vibrio* TIM; altering this to serine (present in all other TIMs) shifts the stability toward the mesophilic enzyme.

69. Wallon G, Lovett ST, Magyar C, Svingor A, Szilagyi A, Závodszky P, Ringe D, Petsko GA: **Sequence and homology model of IPMDH from the psychrophilic bacterium *Vibrio sp.15* suggest reasons for thermal instability.** *Protein Eng* 1997, 10:665-672.

70. Gerike U, Danson MJ, Russell NJ, Hough DW: **Sequencing and expression of the gene encoding a cold-active citrate synthase from an Antarctic bacterium, strain DS2-3R.** *Eur J Biochem* 1997, 248:49-57.

71. Querol E, Perez-Pons JA, Mozo-Villarias A: **Analysis of protein conformational characteristics related to thermostability.** *Protein Eng* 1996, 9:265-271.

72. Wrba A, Schwaiger A, Schultes V, Jaenicke R, Závodszky P: **GAPDH from the extreme thermophilic eubacterium *Thermotoga maritima*.** *Biochemistry* 1990, 29:7585-7592.

73. Závodszky P, Kardos J, Svingor A, Petsko GA: **Adjustment of conformational flexibility is a key event in the thermal adaptation of proteins.** *Proc Natl Acad Sci USA* 1998, 95:7406-7411.

74. Dym O, Mevarech M, Sussman JL: **Structural features that stabilize halophilic MDH from an archaebacterium.** *Science* 1995, 267:1344-1346.

75. Eisenberg H: **Life in unusual environments: progress in understanding the structure and function of enzymes from extreme halophilic bacteria.** *Arch Biochem Biophys* 1995, **318**:1-5.

76. Frolow F, Harel M, Sussman JL, Mevarech M, Shoham M: **Insights into protein adaptation to a saturated salt environment from the crystal structure of a halophilic 2Fe–2S ferredoxin.** *Nat Struct Biol* 1996, **3**:452-458.

77. Madern D, Pfister C, Zacchai G: **Mutation at a single acidic amino acid enhances the halophilic behaviour of malate dehydrogenase from *Haloarcula marismortui* in physiological salts.** *Eur J Biochem* 1995, **230**:1088-1095.

78. • Britton KL, Stillman TJ, Yip KSP, Forterre P, Engel PC, Rice DW: **Insights into the molecular basis of salt tolerance from the study of glutamate dehydrogenase from *Halobacterium salinarium*.** *J Biol Chem* 1998, **273**:9023-9030.
Comparative modeling of HsGluDH based on GluDH from *Pyrococcus furiosus* (47% sequence identity) with other related GluDH structures as a guide for small insertions or deletions, was used to provide insight into the molecular basis of halophilic adaptation. Features that may contribute to the stability of HsGluDH in high salt include decoration of the surface with acidic residues and a reduction in the number of surface-exposed lysine residues.

79. • Pieper U, Kapadia G, Mevarech M, Herzberg O: **Structural features of halophilicity derived from the crystal structure of DHFR from the Dead Sea halophilic archaeon, *Haloferax volcanii*.** *Structure* 1998, **6**:75-88.
The crystal structure of dihydrofolate reductase (DHFR) at 2.6 Å resolution reveals the well-known overall fold of other DHFRs. The moderately halophilic nature of the enzyme, which retains its native state at monovalent salt concentrations as low as 0.5 M, is characterized by a lack of the structural features that are known to occur in extremely halophilic proteins. Differences in the percentage of carboxylate groups are statistically insignificant; similarly, no significant increase in ion-pair interactions is detectable.

80. • Jolley KA, Russell JM, Hough DW, Danson MJ: **Site-directed mutagenesis and halophilicity of dihydrolipoamide dehydrogenase from the halophilic archaeon, *Haloferax volcanii*.** *Eur J Biochem* 1997, **248**:362-368.
A homology-modeled structure, based on the crystal structure of dihydrolipoamide dehydrogenase from *P. fluorescence*, allowed the prediction of a putative K^+-binding site at the dimer interface. This binding site seems to be involved in the stabilization of the enzyme at high salt and high temperature. Similar local effects have been reported for malate dehydrogenase from *H. marismortui* [74], supporting the view that the structural basis of halophilicity can only be understood at the level of the precise three-dimensional structure.

81. Kiefhaber T, Schmid FX, Renner M, Hinz H-J, Hahn U: **Stability of recombinant Lys25-RNase T1.** *Biochemistry* 1990, **29**:8250-8257.

82. Blöchl E, Rachel R, Burggraf S, Hafenbradl D, Jannasch HW, Stetter KO: ***Pyrolobus fumarii* represents a novel group of archaea, extending the upper temperature limit for life to 113°C.** *Extremophiles* 1997, **1**:14-21.

83. van den Burg B, Vriend G, Veltman OR, Venema G, Eijsink VGH: **Engineering an enzyme to resist boiling.** *Proc Natl Acad Sci USA* 1998, **95**:2056-2060.

84. de Filippis V, de Antoni F, Frigo M, Polverino de Loreto P, Fontana A: **Enhanced protein thermostability by Ala→Aib replacement.** *Biochemistry* 1998, **37**:1686-1696.

85. Mendel D, Cornish VW, Schultz PG: **Site-directed mutagenesis with an expanded genetic code.** *Annu Rev Biophys Biomol Struct* 1995, **24**:435-462.

86. Veltman OR, Vriend G, Berendsen HJC, van den Burg B, Venema G, Eijsink VGH: **A single Ca^{2+} binding site is crucial for the Ca^{2+}-dependent thermal stability of thermolysin-like proteases.** *Biochemistry* 1998, **37**:5312-5319.

87. Noll KM, Vargas M: **Recent advances in genetic analyses of hyperthermophilic archaea and bacteria.** *Arch Microbiol* 1997, **168**:73-80.

88. •• Russell RJM, Ferguson JMC, Hough DW, Danson MJ, Taylor GL: **The crystal structure of citrate synthetase from the hyperthermophilic archaeon *Pyrococcus furiosus* at 1.9 Å resolution.** *Biochemistry* 1997, **36**:9983-9994.
Comparing the *Pyrococcus* citrate synthetase (CS) enzyme with CS from pig and *Thermoplasma acidophilum*, all three share a similar fold and the same acid-base catalysis. Given these similarities, trends in going from mesophilic CS to hyperthermophilic CS are detectable. These include increased compactness, more intimate subunit interactions, an increase in clustered intersubunit ion pairs, the shortening of loops and a reduction in the number of thermolabile residues.

89. Macedo-Ribeiro S, Darimont B, Sterner R, Huber R: **Small structural changes account for the high thermostability of 1[4Fe–4S] ferredoxin from *Thermotoga maritima*.** *Structure* 1996, **4**:1291-1301.

90. • Pfeil W, Gesierich U, Kleemann GR, Sterner R: **Ferredoxin from the hyper-thermophile *Thermotoga maritima* is stable beyond the boiling point of water.** *J Mol Biol* 1997, **272**:591-596.
The 60 residue *T. maritima* ferredoxin contains a single [4Fe–4S] cluster that is irreversibly released upon unfolding. Since there are no further heat-absorbing processes, the reaction can be interpreted thermodynamically. Polar interactions contribute significantly to the stability. This shows its maximum at 45°C, when ΔG amounts to 39 kJ/mol. Thus, the protein is thermodynamically no more stable than average mesophilic proteins. Due to its small size and the low heat capacity change ΔC_p of unfolding (3.6 kJ/molK), the ΔG versus temperature profile is extremely shallow, thus explaining the wide range of stability between −20 and +125°C.

91. Hatanaka H, Tanimura R, Katoh S, Inagaki F: **Solution structure of ferredoxin from the thermophilic cyanobacterium *Synechococcus elongatus* and its thermostability.** *J Mol Biol* 1997, **268**:922-933.

92. Yip KSP, Stillman TJ, Britton KL, Artymiuk PJ, Baker PJ, Sedelnikova SE, Engel PC, Pasquo A, Chiaraluce R, Consalvi V *et al.*: **The structure of *Pyrococcus furiosus* GluDH reveals a key role for ion-pair networks in maintaining enzyme stability at extreme temperatures.** *Structure* 1995, **3**:1147-1158.

93. • Yip KSP, Britton KL, Stillman TJ, Lebbink J, de Vos WM, Robb FT, Vetriani C, Maeder D, Rice DW: **Insights into the molecular basis of thermal stability from the analysis of ion-pair networks in the GluDH family.** *Eur J Biol* 1998, **255**:336-346.
A comparison of the crystal structures of glutamate dehydrogenases (GluDHs) from *P. furiosus*, *Clostridium symbiosum* and *E. coli* gives clear evidence for the formation of extended networks of ion pairs being a significant contribution to thermal stability. Homology modeling indicates that a decrease in temperature stability correlates with increased 'fragmentation' of ion-pair networks.

94. Komdörfer I, Steipe B, Huber R, Tomschy A, Jaenicke R: **The crystal structure of holo-GAPDH from the hyperthermophilic bacterium *Thermotoga maritima* at 2.5 Å resolution.** *J Mol Biol* 1995, **246**:512-521.

95. Hennig M, Darimont B, Sterner R, Kirschner K, Jansonius JN: **2.0 Å structure of indole-3-glycerol phosphate dehydrogenase from the hyperthermophile *Sulfolobus solfataricus*: possible determinants of protein stability.** *Structure* 1995, **3**:1295-1306.

96. •• Wallon G, Kryger G, Lovett ST, Oshima T, Ringe D, Petsko GA: **Crystal structures of *E. coli* and *S. typhimurium* 3-isopropyl MDH and comparison with their thermophilic counterpart from *Thermus thermophilus*.** *J Mol Biol* 1997, **266**:1016-1031.
All three 3-isopropyl malate dehydrogenase (IPMDH) structures show a resolution < 2.1 Å. The main stabilizing features in thermophilic IPMDH are an increased number of ion pairs, additional hydrogen bonds, a proportionately larger and more hydrophobic subunit interface, shortened N and C termini, and a larger number of proline residues. Aromatic–aromatic interactions, helical capping and cavities seem to be of lesser importance. The overall flexibility (monitored by the B factors of equivalent residues) is reduced in the case of *S. typhimurium* IPMDH.

97. • Akanuma S, Yamagishi A, Tanaka N, Oshima T: **Serial increase in the thermal stability of IPMDH from *Bacillus subtilis* by experimental evolution.** *Protein Sci* 1998, **7**:678-705.
The mesophilic IPMDH from *B. subtilis* was expressed in an IPMDH-deficient mutant of *T. thermophilus*. Variants growing at temperatures higher than the stability limit of wildtype *B. subtilis* IPMDH (56°C) were selected at 61–70°C. Three point mutants, I95L, M292I and T308I, and double and triple mutants thereof were isolated and characterized. The stability of the variants was improved, showing additivity in their stabilization effects, and the triple mutant had also higher specific activity than the wildtype protein.

98. • Aoshima M, Oshima T: **Stabilization of *E. coli* IPMDH by single amino acid substitutions.** *Protein Eng* 1997, **10**:249-254.
In attempting to define the key positions responsible for the high stability of IPMDH from *Thermus thermophilus* (Tt) and *Thermus aquaticus* (Ta), sequence comparisons of all 20 known IPMDHs were carried out. As a result, a four-residue motif, which is present in the Tt and Ta sequences only, was found located between two highly conserved stretches of the sequence. Introducing this motif into *E. coli* IPMDH, one by one, enhanced thermostability, whereas introducing the whole motif caused destabilization. The three-dimensional structure allows a tentative explanation of the mechanism.

THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 273, No. 15, Issue of April 10, pp. 9023–9030, 1998
Printed in U.S.A.

Insights into the Molecular Basis of Salt Tolerance from the Study of Glutamate Dehydrogenase from *Halobacterium salinarum**

(Received for publication, October 16, 1997, and in revised form, December 3, 1997)

K. Linda Britton‡, Timothy J. Stillman‡, Kitty S. P. Yip‡, Patrick Forterre§, Paul C. Engel¶, and David W. Rice‡‖

From the ‡The Krebs Institute for Biomolecular Research, Department of Molecular Biology and Biotechnology, University of Sheffield, Sheffield S10 2TN, United Kingdom, §Institute Genetique et Microbiologie, Bâtiment 409, Université Paris-Sud, 91405, Orsay Cedex, France, and ¶Department of Biochemistry, University College Dublin, Belfield, Dublin 4, Ireland

A homology-based modeling study on the extremely halophilic glutamate dehydrogenase from *Halobacterium salinarum* has been used to provide insights into the molecular basis of salt tolerance. The modeling reveals two significant differences in the characteristics of the surface of the halophilic enzyme that may contribute to its stability in high salt. The first of these is that the surface is decorated with acidic residues, a feature previously seen in structures of halophilic enzymes. The second is that the surface displays a significant reduction in exposed hydrophobic character. The latter arises not from a loss of surface-exposed hydrophobic residues, as has previously been proposed, but from a reduction in surface-exposed lysine residues. This is the first report of such an observation.

In highly saline environments, for example in salt lakes or in desiccating salt marshes, where salt concentrations can exceed 3 M, the dominant microorganisms are extremely halophilic Archaea (1). These halophilic organisms accumulate inorganic ions within the cell at concentrations equivalent to or greater than that of the environment (2), and their proteins are therefore specialized to function under high salt conditions. The ease with which these organisms are grown and the absence of a necessity for aseptic conditions makes them very attractive for commercial applications including, among others, production of bio-degradable plastics (3) and cosmetics (4). Furthermore, salt, like solvents, dehydrates enzymes, and therefore, information about the survival mechanisms of halophiles could well enable other enzymes to be modified to function efficiently in other solvents more relevant to the conditions used in many industrial processes (5).

The amino acid sequence of glutamate dehydrogenase (GluDH)[1] from *Halobacterium salinarum* (*Hs*) reveals that this enzyme contains a high number of acidic amino acids (6, 7), but as yet, a three-dimensional structure is not available for this GluDH. In previous work we have determined the high resolution structures of the GluDHs from the mesophile *Clostridium symbiosum* (*Cs*) (8, 9) and from the hyperthermophilic archaeon *Pyrococcus furiosus* (*Pf*) (10). The close similarity in the sequence of the halophilic enzyme to these other GluDHs implies that the proteins possess closely related three-dimensional structures. This led us to carry out a homology-based modeling study on this halophilic enzyme, enabling us to examine the distribution of particular residues and thereby contributing to a further understanding of the molecular basis of salt tolerance.

EXPERIMENTAL PROCEDURES

Sequence Alignment—The GluDH sequences from a diverse range of species show a very high degree of sequence similarity (6, 11, 12). Furthermore, the structures of this enzyme from *Cs* and *Pf* are also very similar (10), strongly suggesting that the core of the three-dimensional structures of GluDH are highly conserved, and therefore these structures can serve as models for all other hexameric GluDHs. The sequences of the GluDHs from *Cs* (11) and *Pf* (13, 14) were aligned against one another using their three-dimensional structures as a guide. The alignment of the sequence of the GluDH from *Hs* (6) against those of the *Cs* and *Pf* enzymes was greatly simplified by its similarity to the latter (47% identity) (Fig. 1). Of the 68 residues strongly conserved across the family of GluDHs, 57 are conserved in the *Hs* enzyme. This suggests that the key residues concerned with the maintenance of the catalytic properties and structural framework of the enzyme are not modified by the necessity of the halophilic enzyme to operate in high salt conditions. Throughout this paper, unless specified, the *Hs* GluDH sequence numbering is used to identify equivalent residues in the other GluDH sequences.

Construction of the Model of Hs GluDH—To date the structures of six glutamate dehydrogenases have been determined; *C. symbiosum* (8) *Escherichia coli*,[2] *Neurospora crassa*,[3] *Pyrococcus furiosus* (10), *Thermotoga maritima* (15), and *Thermococcus litoralis*.[4] The r.m.s. fit based on 436 residues of the *Cs* structure to that of *E. coli* is approximately 0.8 Å despite a sequence identity of only 52%. Even between the *Cs* and *Pf* enzymes (identity 32%), the r.m.s. fit based on 210 out of 419 equivalenced residues is only 1.0 Å. Thus, given the close sequence relationship between the *Hs* and *Pf* enzymes (identity 47%), we have chosen to model the *Hs* enzyme onto the *Pf* structure using the other GluDH structures as a guide where small insertions or deletions make these more appropriate.

The sequence information from the alignment together with the program FRODO (16) was used to produce an atomic model for the halobacterial GluDH based primarily on the structure of the more closely related *Pf* enzyme. The major differences between this model and the structure of the *Pf* enzyme relate to differences in the path of the main chain caused by the occurrence of small insertions and deletions. Such differences are found in six regions and involve only 35 residues out of the 428 residues present in the model. First, at the N terminus, the level of sequence homology between *Pf* and *Hs* GluDH is particularly poor, and an additional 21 residues are found in the latter. In comparison, the clostridial enzyme has an additional 15 residues with respect to pyrococcal GluDH, with these residues folding to form

* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‖ To whom correspondence should be addressed. Tel.: 44 114 2224242; Fax: 44 114 2728697; E-mail: D.Rice@Sheffield.ac.uk.

[1] The abbreviations used are: GluDH, glutamate dehydrogenase; *Hs*, *Halobacterium salinarum*; *Cs*, *Clostridium symbiosum*; *Pf*, *Pyrococcus furiosus*; r.m.s., root mean square; e, electron.

[2] T. J. Stillman, I. S. B. Abeysinghe, J. L. Dean, P. C. Engel, and D. W. Rice, manuscript in preparation.

[3] T. J. Stillman, D. W. Rice, A. M. Fuentes, and I. Connerton, manuscript in preparation.

[4] K. S. P. Yip, T. J. Stillman, F. Rabb, and D. W. Rice, manuscript in preparation.

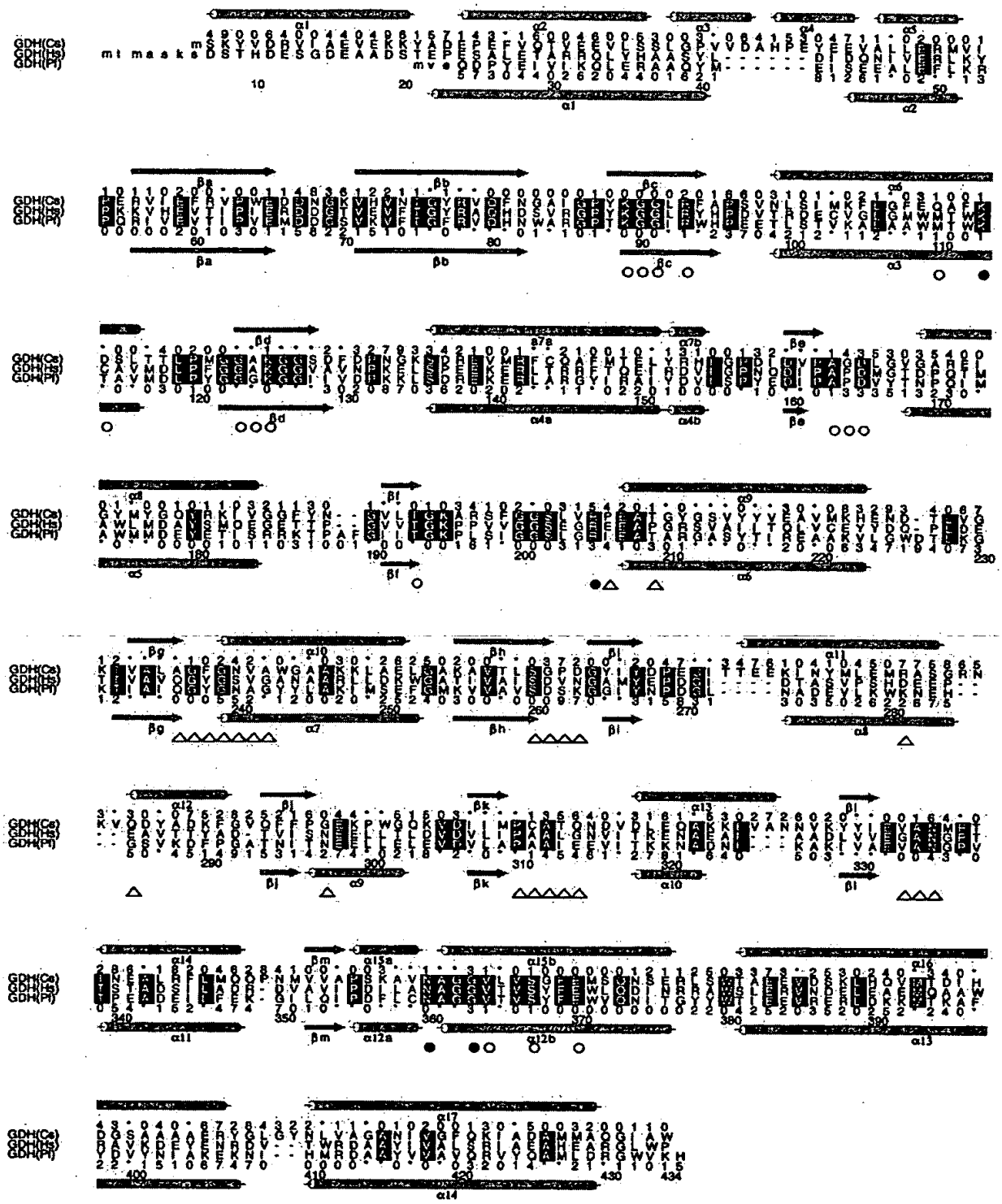


FIG. 1. **Alignment of the sequences of the GluDHs (*GDH*) from *Cs*, *Hs*, and *Pf*.** The amino acids for each sequence are identified by the single letter code, with those in *lower case* representing the N-terminal amino acids that are disordered in the x-ray structures of *Cs* and *Pf* GluDH and where, in the *Hs* enzyme, there is no equivalent in either of the other two GluDH sequences. These residues have not been included in the analysis. Positions of insertions and deletions between the sequences are indicated by −. The secondary structural elements of the three-dimensional structure of the *Cs* GluDH are shown above, with helices represented by *cylinders* and strands shown as *arrows*. A similar representation of the three-dimensional structure of the *Pf* GluDH is shown below. The *numbering* displayed beneath the alignment corresponds to the halophilic enzyme sequence. Residues that are located within 6 Å of any atom of the bound glutamate substrate or the dinucleotide cofactor are indicated by O and △, respectively. Where residues lie within this limit for both bound substrate and cofactor, the residues are highlighted by ●. The solvent-accessible surface area of the side chains in the hexamer of the three GluDHs were calculated by the method of Lee and Richards

an extra helix. We have therefore chosen to ignore the first six of the additional residues in the halophilic enzyme and to model the remainder as a helix, as found in clostridial GluDH. Inspection of the model appears to support this, as all six acidic residues in this region fall on an exposed face of the helix, and the only other charged residue (His-11) is partially exposed to solvent.

Elsewhere, the *Cs* GluDH structure was only used to model those loop regions where differences in length in the alignment occurred between the two archaeal enzymes but where similarities in length with the clostridial enzyme were noted. This affects the structure around residues 186–189, 226–228, and 250–253 inclusively in the *Hs* GluDH. The construction of this hybrid model therefore involved some local rebuilding and geometry regularization at the "annealing" points between the two structures but also at one other location where there was a deletion of one residue in the *Hs* sequence with respect to both the *Cs* and *Pf* GluDH structures (between residues 292 and 293 in *Hs* GluDH). Finally, the *Hs* sequence is one residue shorter at the C terminus.

This model was then used to provide the structural backbone onto which the relevant side-chain residues were substituted to produce a model for the *Hs* GluDH. If side chains were in common between the GluDH structure being used (primarily that of the *Pf* enzyme) and the *Hs* GluDH sequence, their conformations were retained. Where side chains required substitution, the position of the replaced residue was maintained as far as possible unless this introduced unacceptable steric clashes. Analysis of the final model using PROCHECK (17) suggested that the torsion angle distribution is typical for that seen in a high resolution protein structure, and there are no residues with disallowed Ramachandran angles.

Analysis of the Model—The solvent-accessible surface areas for each atom of both a monomer and the hexamer of the three enzymes were calculated using the algorithm of Lee and Richards (18), excluding the solvent molecules of the models. The resulting solvent-accessible areas for each residue were expressed as a fraction of the total solvent-accessible surface area for each type of amino acid (19). Atoms that recorded different solvent-accessible areas between the monomer and the hexamer were defined as the buried surface area on hexamer assembly. The definitions of Miller *et al.* (20) for nonpolar, polar, and charged constituents of proteins were used to tabulate the chemical composition of the surface.

RESULTS

Differences in Amino Acid Composition

The comparison of the amino acid compositions between halophilic proteins and their mesophilic counterparts has highlighted the emergence of three general trends in halophilic proteins: an excess of acidic over basic residues, an increase in the "borderline" hydrophobic residues serine and threonine, and a collective decrease in the strongly hydrophobic residues valine, isoleucine, leucine, and phenylalanine (21). Analysis of the amino acid compositions for the mesophilic, halophilic, and hyperthermophilic GluDHs (Table I) shows that the halophilic GluDH contains significantly more acidic residues, with 64% of the total number of charged residues being either aspartate or glutamate compared with 51 and 53% of such residues in the *Pf* and *Cs* enzymes, respectively. In total, the halophilic enzyme contains 77 acidic and 44 basic residues in each subunit, which gives rise to an overall negative charge of 198 for the hexamer. This compares with the significantly lower values for the net charge of −18 and −42 for the hexamers of *Pf* and *Cs* GluDH, respectively.

Although the number of glutamate, histidine, and arginine residues is similar for all three enzymes, the number of aspartate residues increases from 24 to 37 between the hyperthermophilic and halophilic GluDHs. At the same time, there is a dramatic reduction in the number of lysines from 32 to 13 in the same two enzymes. Similar increases in acidic residues and lowered lysine content have been found in a comparison between the sequences of elongation factor EF-Tu from *Halobacterium marismortui* and its counterpart from the mesophile *Methanococcus vannielii* (22). Consideration of the other potentially significant differences in composition shows there to be an increase in the proportion of threonine and a decrease in the proportion of phenylalanine in the halophilic GluDH compared with both enzymes and a reduction in isoleucine and increase in serine with respect to the *Pf* enzyme, all of which are fully consistent with the trends noted by Lanyi (21).

TABLE I
Amino acid composition of the GluDHs from Cs (11), Hs (6) and Pf (13, 14)

Amino Acid	*Cs*	*Hs*	*Pf*
Asp	24	37	24
Asn	20	13	15
Ala	43	49	39
Arg	18	24	20
Val	45	39	35
Pro	20	24	19
Gly	48	33	34
Glu	34	40	36
Gln	16	12	11
Cys	2	3	1
Met	17	10	12
Ile	20	22	35
Leu	27	27	24
Lys	27	13	32
His	6	7	5
Tyr	21	15	20
Phe	19	4	10
Trp	5	9	10
Ser	20	22	14
Thr	17	31	23
Total	449	434	419

Location of the Sequence Differences on the Three-dimensional Structure of GluDH

Analysis of Changes in the Buried Core—The subunit structure of the *Pf* GluDH is shown schematically in Fig. 2. Each subunit in this hexameric enzyme is organized into two domains separated by a deep cleft, which forms the active site (Fig. 2). Sequence substitutions of totally buried residues almost exclusively involves conservative replacements from within the set of hydrophobic amino acids (Fig. 1). Substitutions within the subset of largely buried strongly hydrophobic residues (defined as having between 0 and 20% surface area accessible to solvent) commonly involves exchanges from within the set of hydrophobic amino acids, although the replacement by borderline hydrophobic residues such as threonine is also observed (Fig. 1). The replacement of hydrophobic residues that are partially exposed to solvent (defined as having at least 20% of the residue surface accessible to solvent) frequently involves modification to a polar or charged amino acid.

Analysis of the Intersubunit Interface—Analysis of the nature of the subunit surface that is buried on assembly of the hexamer (Table II) shows that the smaller proportion of nonpolar constituents observed for the solvent-accessible surface of the halophilic enzyme is repeated at this interface. Not surprisingly and in contrast to the solvent-accessible surface, the

(18). The resulting accessible areas for each side chain were then expressed as a fraction of the total solvent-accessible surface area for each type of amino acid (19). The exposed surface area for the residues in *Cs* and *Pf* GluDHs are given as a sequence of integers from 0 to 9 immediately above or below the sequences for the *Cs* and *Pf* enzymes, respectively, where each *number* represents residues whose side chains have 1–10, 11–20, 21–30%, etc. of their surface solvent-accessible. Totally buried residues are indicated by a *. References for the sequences given are *Cs* (11), *Hs* (6), and *Pf* (13, 14).

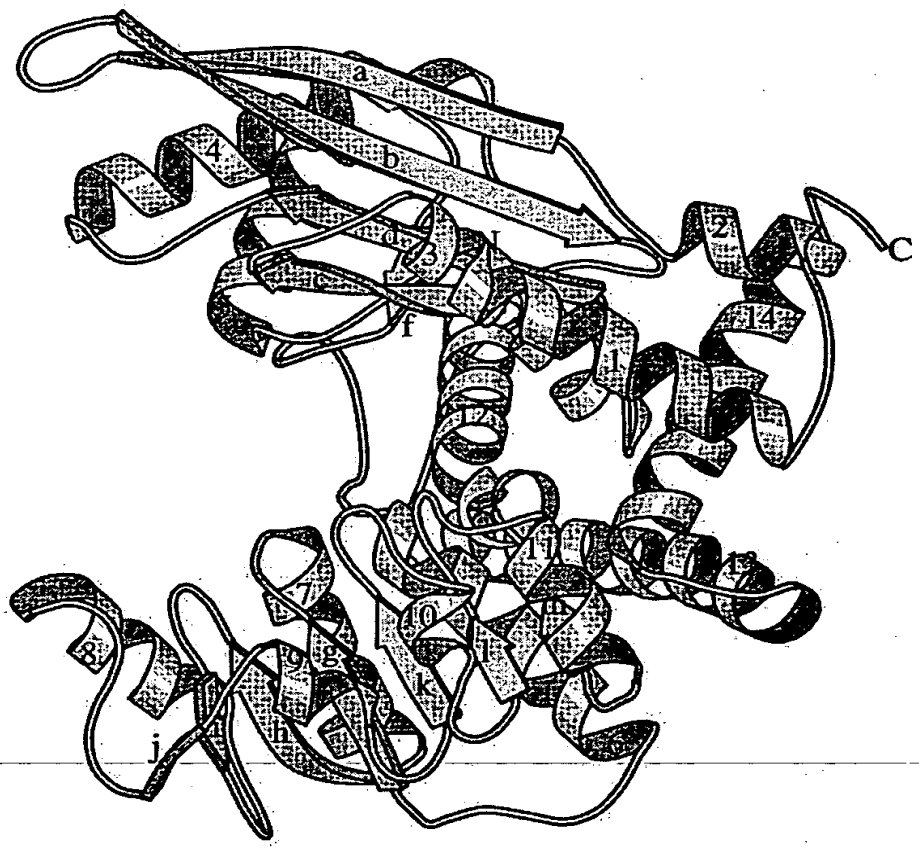


FIG. 2. **A MOLSCRIPT representation of the fold of the polypeptide chain of a single subunit of the *Pf* GluDH is shown schematically (27).** In this view, the 3-fold axis of the GluDH hexamer runs vertically. The enzyme is organized into two domains separated by a deep cleft. Domain I (upper) is responsible for subunit assembly, whereas domain II (lower) contains the nucleotide binding site. The elements of secondary structure are labeled with *numbers* representing helices (1–14) and *letters* for strands (a-m).

TABLE II

The overall solvent-accessible surface areas for the GluDHs from Cs, Hs, and Pf calculated by the method of Lee and Richards (18) for both the monomers and hexamers

Analysis of the chemical composition of the solvent-accessible surface area and the surface areas that are buried at the subunit interfaces in the hexameric GluDHs from the *Cs*, *Hs*, and *Pf*, respectively, in terms of nonpolar, polar, and charged constituent atoms. These calculations are based on the method of Miller *et al.* (20).

	Cs	*Hs*	*Pf*
Solvent-accessible surface areas (Å^2)			
Monomer	18,500	18,500	17,000
Hexamer	88,000	93,000	82,500
Characteristics of the solvent-accessible region (%)			
Nonpolar	47	43	48
Polar	29	25	24
Negatively charged	13	24	17
Positively charged	11	8	11
Characteristics of the subunit interface region (%)			
Nonpolar	51	44	48
Polar	26	28	25
Negatively charged	7	11	10
Positively charged	16	17	17

number of charged residues and the charge balance at the intersubunit interface for these three GluDHs is very similar (Table II). This similarity is not surprising but strongly suggests that there are no gross errors associated with the modeling of the *Hs* enzyme.

The Nature of the Solvent-accessible Surface Area—Characteristics of the surface accessible to the solvent in the hexamers of the hyperthermophilic and mesophilic GluDHs from *Pf* and *Cs* have been analyzed previously (10). This study showed that although the proportion of the nonpolar surface was constant between these two enzymes, there was an increase in the charged nature of the hexamer surface in the hyperthermophilic *Pf* enzyme (24 and 28% of the solvent-accessible area in the *Cs* and *Pf* enzymes, respectively) and a corresponding decrease in the occurrence of polar groups relative to the *Cs* GluDH. Extending the study to include the *Hs* GluDH model (Table II) revealed a smaller proportion of nonpolar components and an even larger proportion of solvent-accessible charged groups relative to the *Cs* and *Pf* enzymes. Further examination of the area of the molecular surface that carries a formal charge has revealed that in each of the three GluDHs, the solvent-accessible surface area contains a higher proportion of negatively charged groups compared with the proportion of positively charged groups. However, this ratio is dramatically increased for the halophilic enzyme model, resulting in a surface predominantly covered in negatively charged residues. This is consistent with results obtained from the structure determinations of the malate dehydrogenase and the ferredoxin from the halophile *Haloarcula marismortui* (23, 24). The net charge density for the hexamer of the halophilic GluDH is -2.6×10^3 eÅ^{-2}, far greater than the values of -0.9×10^3 eÅ^{-2} and -0.5×10^3 eÅ^{-2} for its mesophilic and hyperthermophilic counterparts, respectively, and comparable to the reported net charge densities of other halophilic proteins (24). These significant differences in the acidic nature of the hexamer surface can be seen both as a function of the electrostatic potential and in terms of the distribution of the charged residues on the protein surface (Fig. 3).

If we consider only those residues that are either partially or wholly accessible to the solvent, it is evident from the comparison of the three GluDHs that the larger total of exposed acidic

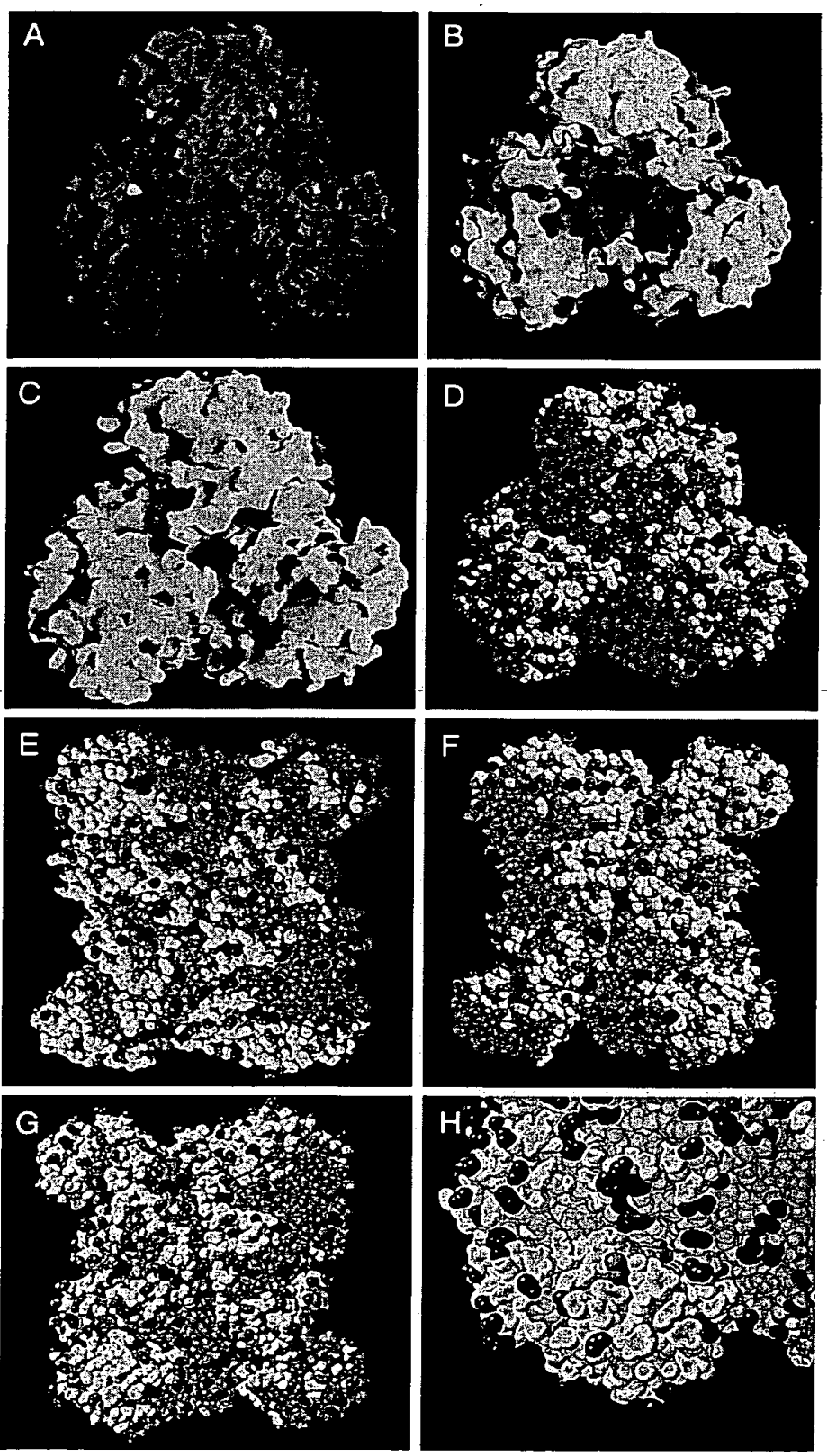


FIG. 3. **Comparative views of the surfaces of the halophilic, mesophilic, and thermophilic GluDHs to illustrate the different surface properties.** *a-c*, views down the 3-fold axis of the molecular surfaces of the *Hs*, *Cs*, and *Pf* enzymes, respectively, to show the electrostatic potential at 0 M salt concentration, prepared using the program GRASP (28) (*red* corresponds to a surface potential less than −20 kcal(mol·electron)$^{-1}$; *blue* corresponds to a potential greater than +20 kcal(mol·electron)$^{-1}$). *d*, the same view of the molecular surface of the hexameric GluDH from *Hs* prepared using the MIDASPLUS program (29, 30), highlighting the distribution of charged residues on the enzyme surface. All guanidinium, imidazole, and amino groups present in arginine, histidine, and lysine residues are shown in *blue*, and carboxyl groups associated with aspartate and glutamate residues are shown in *red*. *e-g*, views parallel to the 2-fold axis (drawn in the same manner as *d*) of the *Cs*, *Pf*, and *Hs* enzymes, respectively. *h*, a close-up of the molecular surface in the active-site region of the *Hs* enzyme. The location of the bound glutamate (*green*) and coenzyme (*magenta*) are shown, revealing the lack of negatively charged residues in the vicinity of the two binding sites, in sharp contrast with the remainder of the protein surface.

TABLE III

Total solvent-accessible surface areas calculated for side chains in the hexamers of the Cs, Hs, and Pf GluDHs, respectively

The data shown focus on those side chains highlighted in the Lanyi (21) survey, which suggests a trend in the halophilic enzymes compared with counterparts from non-halophilic species. Two sets of calculations are included. In the first, the side chain is defined as all atoms of a given residue except backbone nitrogen, carbon, and oxygen atoms. Second, the total surface-accessible surface areas calculated for the hydrophobic constituent atoms for the selected side chains only are also shown. For the latter calculations, all carbon atoms associated with the side chains of each residue type are included.

Side chain	Total solvent-accessible area			Total hydrophobic solvent-accessible area		
	Cs	*Hs*	*Pf*	*Cs*	*Hs*	*Pf*
		Å²			Å²	
Asp	5900	13,700	6500	1800	3500	1800
Glu	12,700	16,600	13,500	4200	4900	4500
Arg	6600	8100	4700	2500	2400	1700
Lys	10,000	3800	14,300	5700	2100	8200
His	1300	1000	1300	900	600	900
Ser	2300	4500	1400	1500	2100	1100
Thr	2700	4500	3500	2000	3000	2300
Phe, Ile, Leu and Val	7200	5300	7500	7200	5300	7500

residues in the halophilic enzyme is dominated by the increase in aspartate. In fact, 36 aspartate residues per subunit fall into this category in the *Hs* enzyme compared with 22 and 23 aspartates in the mesophilic and hyperthermophilic counterparts. An equally dramatic reduction of solvent-accessible lysines is also observed with only 11 lysines of this type in the *Hs* GluDH model compared with 27 and 32 in the structure of the *Cs* and *Pf* enzymes, respectively.

We have also analyzed the solvent-accessible areas of the side-chain components of those residues for which significant differences in composition have been noted. The solvent-accessible area for aspartate in the *Hs* enzyme is more than double that found in the mesophilic and hyperthermophilic counterparts (Table III), reflecting the significant increase in composition associated with this residue type. Similarly, the solvent-accessible area associated with serine and threonine side chains is also greater in the model of *Hs* GluDH. Thus, 29 of the 31 threonines have some part of the residue solvent-accessible in the *Hs* GluDH model compared with 15 of the 17 and 18 of the 23 threonines of the *Cs* and *Pf* GluDH structures, respectively, implying that the increase in threonine residues is concentrated at the protein surface of this enzyme. Likewise, consideration of the serine residues present in each of the three GluDHs shows there to be 20 of the 22, 17 of the 20 and 12 of the 14 accessible to the solvent in the *Hs*, *Cs*, and *Pf* enzymes, respectively. Analysis of the *Hs* GluDH model places many of these surface-accessible serine and threonine residues adjacent to exposed acidic groups. Inspection of the model of the *Hs* enzyme suggests that many of these residues occupy positions where they could potentially hydrogen bond to nearby carboxyl groups.

Examination of the characteristics of the solvent-accessible area shows that there is a somewhat lower proportion of exposed hydrophobic groups in the model of the halophilic GluDH, with 43, 47, and 48% of the surface being hydrophobic in the *Hs*, *Cs*, and *Pf* enzymes, respectively (Table II). Analysis of the contribution for a number of different residue types to the hydrophobic-accessible surface area is presented in Table III. This shows that in the solvent-accessible hydrophobic surface area of the *Hs* model, a reduction of some 2000 Å² is due to changes associated with the subset of strongly hydrophobic amino acid residues (Phe, Ile, Leu, and Val). Moreover, this reduction is offset by the increase in hydrophobic surface brought about by the additional aspartic acid residues. Of all the other changes, one stands out as being predominant. This involves a reduction of 3600 Å² and 6100 Å² in the exposed hydrophobic surface of the halophilic enzyme compared with the *Cs* and *Pf* enzymes, respectively, due to the reduction in the number of exposed lysine residues and the loss of the associated exposed alkyl component of these side chains. Moreover, of the lysines in the *Hs* model, the average solvent-accessible hydrophobic surface per residue is somewhat lower than for the other two GluDHs (175, 210, and 260 Å² for the *Hs*, *Cs*, and *Pf* enzymes, respectively). The explanation for this lies in the fact that those lysines that are retained in the *Hs* enzyme form the subset of such residues that are involved in the functional properties of GluDH and are in fact strongly buried in the structure (defined as a residue with less than 20% solvent-accessible surface). Thus, although there are 18 and 25 exposed lysines in the *Cs* and *Pf* enzymes, there are only 5 in the halophilic counterpart. In contrast, the number of strongly buried lysines is more constant across the three enzymes (8, 9, and 7 for the *Hs*, *Cs*, and *Pf* GluDHs, respectively).

To our knowledge the reduction in hydrophobic surface associated with the depletion of lysine in halophilic enzyme has not been reported even for those comparisons between structures of equivalent halophilic and mesophilic enzymes. Therefore to extend this analysis we have also compared the recently determined structure of the halophilic malate dehydrogenase (23) with the counterpart lactate dehydrogenase from dogfish (Protein Data Bank depositions 1HLP and 6LDH, respectively). Comparison of the amino acid compositions of these two enzymes also shows a lower number of lysines in the halophile (29 and 8 lysine residues in the dogfish and halophilic enzymes, respectively). Furthermore, solvent-accessible surface area calculations revealed a lower surface contribution of nonpolar components in the tetrameric halophilic malate dehydrogenase (42% compared with 50% in the mesophilic counterpart). Analysis of the contributions to this surface from the subset of strongly hydrophobic residues showed a fall of 900 Å² in the halophilic enzyme, insufficient to account for the overall reduction in hydrophobic solvent-accessible surface (5200 Å²). However, the lower lysine content again results in a loss of 5100 Å² in hydrophobic surface, which accounts for the difference between the two enzymes. Furthermore, the reduction in lysine content reported for *H. marismortui* elongation factor EF-Tu compared with the counterpart from the mesophile *M. vannielii* (22) is also consistent with our observations.

d) Analysis of the Active Site—The active site of this enzyme family has been located following analysis of the binary complexes of *Cs* GluDH with NAD^+ and glutamate (8, 9)[5] and of the *Pf* enzyme with NAP^+. Analysis of the region around the glutamate binding pocket has shown that there are 20 amino acid residues that have at least 1 atom lying within 6 Å of any atom of the glutamate substrate (Fig. 1). Of these 20 residues, 16 are completely conserved across all three enzymes. At three of the remaining four positions (110, 161, and 164) the residues of the hyperthermophilic and halophilic enzymes are identical, and the final difference involves a substitution at residue 114 to cysteine in *Hs* GluDH, again a substitution that can be found in other members of the GluDH family (11). Overall, therefore, it would appear that the halophilic enzyme is remarkably similar to its mesophilic and hyperthermophilic counterparts in the region of the active site. Interestingly, we note that the assignment of the region of the enzyme surface that forms the active site in *Hs* GluDH would have been a potentially simple procedure even in the absence of direct structural information,

[5] K. L. Britton, T. J. Stillman, K. S. P. Yip, and D. W. Rice, manuscript in preparation.

ferredoxin is indeed adapted to the intracellular high-salt environment.

MATERIALS AND METHODS

Chemicals

Sepharose-4B, diethyl aminoethyl-cellulose (DEAE-cellulose), acrylamide, and 3-(*N*-morpholino)propane-sulfonic acid (MOPS) were procured from Sigma (St. Louis, MO). Bacteriological peptone (L37) was from Oxoid (Hampshire, UK). All other chemicals were of analytical grade and were used as supplied. Double glass-distilled water was used for purification and for all experiments.

Bacterial strain and growth medium

The *H. salinarum* (strain M1) cells were grown in a complex medium containing 250 g NaCl, 20 g $MgSO_4 \cdot 7H_2O$, 2 g KCl, 3 g sodium citrate, and 10 g bacteriological peptone per liter. The inoculum was built up by at least three successive transfers in the fresh medium. The autoclaved medium was inoculated with 5% (v/v) of this inoculum and incubated in a rotary shaker at 37°C and 150 rpm. For frequent use the culture was maintained by transfer to fresh medium at 4-day intervals. When not required for prolonged periods it was maintained by subculturing, once every 20 days, on slants prepared by solidifying the above medium with 1.8% (w/v) agar.

Buffer solutions

The following buffers were used during various stages of purification: 0.025 M MOPS + 4.5 M NaCl, pH 7.5 (buffer A); 40% $(NH_4)_2SO_4$ + 0.025 M MOPS, pH 7.5 (buffer B); 60% $(NH_4)_2SO_4$ + 0.025 M MOPS, pH 7.5 (buffer C); 30% $(NH_4)_2SO_4$ + 0.025 M MOPS, pH 7.5 (buffer D). Circular dichroism (CD), fluorescence, and optical absorption studies were performed in 10 mM sodium phosphate buffer (pH 7.3) supplemented with various concentrations of NaCl as noted for specific experiments.

Purification of ferredoxin

H. salinarum cells were grown to stationary growth phase ($A_{660} \sim 1.5$) in the complex medium. The cells were harvested by centrifugation at 2000 × *g* for 30 min and resuspended in 2 ml of buffer A/g of cells. The resuspended cells were sonicated in ice for 5 × 1 min. The sonicate was frozen and thawed once and then subjected again to sonication for a second round (5 × 1 min). The sonicate was then centrifuged (Beckman L8–60M) at 50,000 × *g* for 2 h at 4°C. The supernatant was collected and diluted with half the volume of buffer A. This was dialyzed overnight against buffer B to exchange the NaCl with 40% $(NH_4)_2SO_4$. The precipitate was removed by centrifugation (50,000 × *g*, 2 h), and the new supernatant was dialyzed again against buffer C. The precipitate appearing at this stage was removed by centrifugation (75,000 × *g*, 1 h). The clear supernatant obtained was loaded onto a Sepharose-4B column (2 cm × 35 cm) preequilibrated with buffer C. Under these conditions HsFd binds to the column. The column was washed with the same buffer until the A_{280} reached <0.1, then elution was carried out with a linear reverse gradient of $(NH_4)_2SO_4$ (60% to 0%) in 25 mM MOPS buffer pH 7.5 (1 ml/min), and absorbance of the eluent was monitored at 420 nm. The first peak fractions eluting at ~26% $(NH_4)_2SO_4$ containing ferredoxin were pooled and were readjusted to 60% with neutralized saturated solution of $(NH_4)_2SO_4$. Some precipitate appearing at this stage was removed by filtration. The clear filtrate was loaded onto a DEAE-cellulose column (1 cm × 20 cm) preequilibrated with buffer C and was washed with 5 bed volumes of the same buffer. Ferredoxin, which bound as a very dark band at the top of the column, was eluted with buffer D. The condensed dark brown band spread over the column at this stage and was eluted over several fractions. The fractions with A_{420}:A_{275} >0.2 were pooled, and an equal volume of saturated, neutralized $(NH_4)_2SO_4$ was added and rechromatographed on a DEAE-cellulose column. This chromatographic step was repeated until a A_{420}:A_{275} ratio of 0.33 was finally obtained in buffer C. The protein was then loaded on a small DEAE-cellulose column, eluted with buffer A and dialyzed against it to remove traces of $(NH_4)_2SO_4$. The protein thus obtained was checked by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE), and it was found that the major protein was ferredoxin, along with a small amount of impurities. This protein was further purified on a Sephadex-G50 column (1 cm × 60 cm) preequilibrated with buffer A. The pooled fractions, exhibiting an A_{420}:A_{275} ratio of 0.35, were then concentrated by ultrafiltration (Amicon), using YM10 membrane. Ferredoxin prepared by this method was essentially homogeneous by SDS-PAGE analysis.

Optical measurements

Absorbance spectra were recorded on Shimadzu and Spectronic-1201 spectrophotometers, using cuvettes of 1 cm path length.

CD spectra

CD spectra were recorded on a Jasco-600 Spectropolarimeter. CD data are reported as mean residue ellipticities with a mean residue molar mass of 110. The path lengths of the cuvettes were 1.0, 0.1, or 0.05 cm, depending on the protein concentration and wavelength region. For each spectrum 5–10 scans were co-added, buffers along with the corresponding salt concentration were subtracted, and the spectra were smoothed using mild smoothing function.

Fluorescence measurements

Fluorescence intensities and emission spectra were measured on a Spex spectrofluorimeter with a cuvette of 1 cm path length. All measurements were made with excitation and emission bandwidths of 3.77 nm and 7.54 nm, respectively, using a single photomultiplier. The excitation wavelength was 295 nm. Emission spectra were recorded between 310 and 400 nm. They were baseline corrected for the corresponding buffer. A measurement of the intensity at wavelengths 360 and 330 nm was made for the calculation of the ratio of I_{360}:I_{330}. This ratio was used for the assessment of the overall protein structure (see text for details).

Protein estimation

Protein was estimated by the method of Lowry et al. (1951), as modified by Peterson (1983). Bovine serum albumin was used as the standard. For some experiments the A_{420} value was also used.

RESULTS

Purification and characterization of ferredoxin

H. salinarum ferredoxin was purified by a typical halophilic approach (Werber and Mevarech, 1978) as detailed in Materials and Methods. Its purity was checked by SDS-PAGE, which showed a single band corresponding to ferredoxin. An A_{420}:A_{275} ratio of 0.35 is routinely obtained. The purified protein yield was estimated by Lowry's modified method (Lowry et al., 1951; Peterson, 1983) as well as by its

since it is the only region on the protein surface not to be dominated by the almost uniform coverage by acidic residues (Fig. 3). This may prove to be a general feature of many halophilic enzymes.

Analysis of Ion Pair Networks

The recently reported structure of the halophilic malate dehydrogenase (23) highlighted the formation of clusters of ion pairs, a feature shared by, although more prominent in, the thermophilic malate dehydrogenase from *Thermus flavus*. This feature is absent from the mesophilic dogfish lactate dehydrogenase and is thought to be related to the superior thermal properties of the *T. flavus* and the halophilic enzymes. The recent structure determination of the GluDH from the hyperthermophile *Pf* and its comparison with its counterpart from the mesophile *Cs* has also highlighted a potential role for ion pairs in the determinants of the thermal stability of this enzyme (10). Examination of the halophilic GluDH model suggests that the dramatic 18-residue ion pair cluster in the *Pf* enzyme, which is located across a region of the interface between pairs of dimers, is only partially retained in the *Hs* model, creating two symmetry-related ion-pair networks comprised of four residues (Fig. 4). These findings are consistent with the work on malate dehydrogenase (23) and may explain the apparently greater thermal stability of the halophilic GluDH compared with its mesophilic counterpart (25).

Halophilic Addition

The analysis of the structure of a halophilic 2Fe-2S ferredoxin (24) introduced the concept of halophilic addition. This describes the additional contribution made to the solvent-accessible surface of acidic residues arising from an insertion of an extra small domain of 33 residues that is rich in carboxylates (14 such residues in total), found near the N terminus of this protein. Compared with the mesophilic *Cs* enzyme, the halophilic GluDH is only six residues longer at the N terminus, with none of these being acidic. However, when compared with the hyperthermophilic enzyme, an N-terminal extension of 21 residues, including 6 acidic residues, can be seen. Although the proportion of acidic residues in the N-terminal region of the *Hs* enzyme (29%) is higher than in the protein as a whole (18%), taken together with the comparison with the mesophilic enzyme, the data from the analysis of GluDH are not strongly supportive of the presence of an additional carboxylate-rich domain.

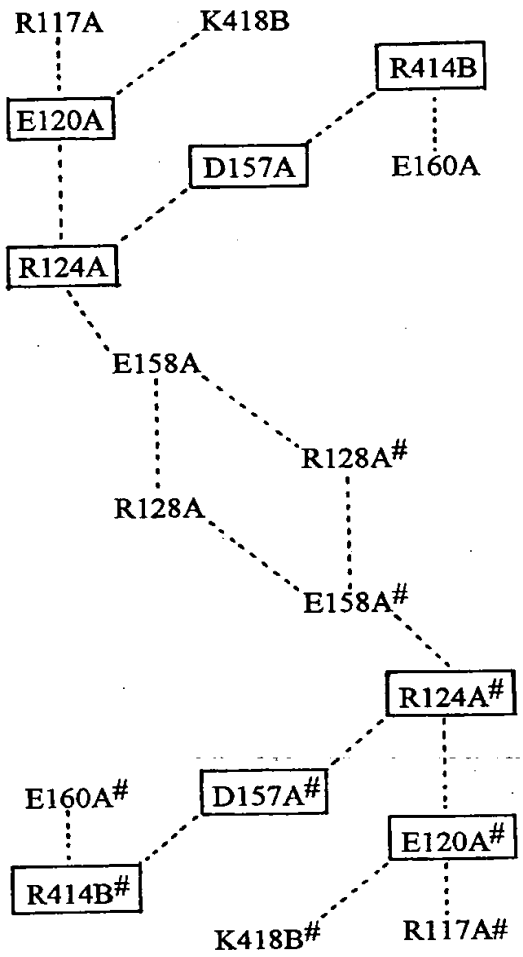


FIG. 4. **A schematic representation of the residues involved in the most extensive ion pair cluster of the *Pf* GluDH involving 18 residues in total across part of the interface region between subunits of the assembled hexamer.** Each amino acid is identified by its *Pf* residue number. The residues highlighted by *boxes* represent those that appear to be maintained in the *Hs* GluDH from the sequence alignment. The *Pf* residues concerned and their *Hs* equivalents are Glu-120 (*E120*) and Glu-141 (*E141*), Arg-124 (*R124*) and Arg-145 (*R145*), Asp-157 (*D157*) and Asp-178 (*D178*), Arg-414 (*R414*) and Arg-430 (*R430*), respectively.

DISCUSSION

Two distinct properties are commonly associated with enzymes from halophiles. The first of these is their ability to catalyze reactions under conditions of extremely high salt. For example, the GluDH from *Hs* is still active in 4 M KCl, whereas that from the mesophile *Cs* is potently inhibited at such high concentrations.[6] The second common property of these halophilic enzymes is that many of those isolated to date are rapidly and irreversibly inactivated by exposure to solutions of low ionic strength. For example, the *Hs* GluDH is markedly unstable below 1 M KCl unless the salt is replaced by a compatible solute such as betaine (26). One immediate question that arises therefore is whether these two properties are related in molecular terms or whether they represent different aspects of the structure/function relationships.

Three features emerge as potentially significant in the comparison of the model of the halophilic enzyme with its nonhalophilic counterparts. The first of these is that the surface of the model of the halophilic GluDH has shown that it is decorated with acidic residues, except in the immediate vicinity of the coenzyme and substrate binding sites. It has been previously argued (23) that the excess surface negative charge is possibly responsible for the formation of a hydration shell that protects the enzyme from aggregation in its highly saline environment. It is not clear to what extent this alone can account for the efficiency with which the halophilic enzymes handle their substrates in the presence of competing high salt. However, it is reasonable to argue that at low ionic strength, the absence of a shielding cloud of counter ions would promote electrostatic repulsion between the closely spaced negative charges on the protein surface, thus destabilizing the protein. Secondly, the analysis of the halophilic GluDH has also highlighted an increased number of surface threonine and serine residues. In the *Hs* model these appear to occupy positions where, were they also to be acidic, they would otherwise result in charge repulsion effects. This feature has been noted in the halophilic ferredoxin from *H. marismortui*, where serine residues have been found on the enzyme surface sandwiched be-

[6] M. Kalinowski and P. C. Engel, unpublished results.

tween two glutamate residues, where the presence of a third carboxylate would be unfavorable (24). Finally, the marked reduction in the number of surface lysine residues is a further dominant feature of the halophilic GluDH and, while helping to increase the overall negative charge on the protein, also serves to decrease the hydrophobic fraction of the solvent-accessible surface. To our knowledge, this is the first report of such an observation. Moreover our analysis of the structure of the halophilic malate dehydrogenase compared with a structure for a mesophilic counterpart strongly supports this finding, showing both a decrease in the number of lysine residues and the consequent marked reduction in the contribution of solvent-accessible hydrophobic surface. At present, the significance of this observation is unclear, but it may be that the presence of significant numbers of alkyl groups on the enzyme surface may well serve to disrupt the production of a well connected hydration shell required in such saline environments, and therefore the long alkyl tails associated with lysines are particularly unfavorable.

For the future, this comparative analysis provides a hypothesis on the molecular basis of salt tolerance that is clearly testable by site-directed mutagenesis. One challenge therefore will be to rationally engineer such properties into mesophilic enzymes to exploit them in an industrial context. To date, this has not been accomplished, and although the structural data on halophilic enzymes may now point the way forward, we should be cautious in assuming that we now understand the structural basis of this phenomena and can manipulate it at will.

Acknowledgments—We thank the United Kingdom Biotechnology and Biological Sciences Research Council, the Wellcome Trust, the European Union Biotechnology Program, and the New Energy and Industrial Development Organization for financial support. The Krebs Institute is a designated Biotechnology and Biological Sciences Research Council Biomolecular Sciences Center.

REFERENCES

1. Woese, C. R., and Fox, G. E. (1977) *Proc. Natl. Acad. Sci. U. S. A.* **74,** 5088–5090
2. Eisenberg, H., Mevarech, M., and Zaccai, G. (1992) *Adv. Protein Chem.* **43,** 1–62
3. Lillo, G. J., and Ridriguez-Valera, F. (1990) *Appl. Environ. Microbiol.* **56,** 2517–2521
4. Galinski, E. A., and Tindall, B. J. (1992) in *Molecular Biology and Biotechnology of Extremophiles* (Herbert, R. D., and Sharp, R. J., eds) pp. 76–114, Blackie and Sons, London
5. Flam, F. (1994) *Science* **265,** 471–472
6. Benachenhou, N., and Baldacci, G. (1991) *Mol. Gen. Genet.* **230,** 345–352
7. Benachenhou-Lahfa, N., Labedan, B., and Forterre, P. (1994) *Gene* **140,** 17–24
8. Baker, P. J., Britton, K. L., Engel, P. C., Farrants, G. W., Lilley, K. S., Rice, D. W., and Stillman, T. J. (1992) *Proteins Struct. Funct. Genet.* **12,** 75–86
9. Stillman, T. J., Baker, P. J., Britton, K. L., and Rice, D. W. (1993) *J. Mol. Biol.* **234,** 1131–1139
10. Yip, K. S. P., Stillman, T. J., Britton, K. L., Artymiuk, P. J., Baker, P. J., Sedelnikova, S. E., Engel, P. C., Pasquo, A., Chiaraluce, R., Consalvi, V., Scandurra, R., and Rice, D. W. (1995) *Structure* **3,** 1147–1158
11. Teller, J. K., Smith, R. M., McPherson, M. J., Engel, P. C., and Guest, J. R. (1992) *Eur. J. Biochem.* **206,** 151–159
12. Britton, K. L., Baker, P. J., Rice, D. W., and Stillman, T. J. (1992) *Eur. J. Biochem.* **209,** 851–859
13. Maras, B., Valiante, S., Chiaraluce, R., Consalvi, V., Politi, L., De Rosa, M., Bossa, F., Scandurra, R., and Barra, D. (1994) *J. Protein Chem.* **13,** 253–259
14. Eggen, R. I. L., Geering, A. C. M., Waldkotter, K., Antranikan, G., and de Vos, W. M. (1993) *Gene* **132,** 143–148
15. Knapp, S., de Vos, W. M., Rice, D., and Ladenstein, R. (1997) *J. Mol. Biol.* **267,** 916–932
16. Jones, T. A. (1985) *Methods Enzymol.* **115,** 157–171
17. Laskovski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallog.* **26,** 283–291
18. Lee, B., and Richards, F. M. (1971) *J. Mol. Biol.* **55,** 379–400
19. Singh, J., and Thornton, J. M. (1992) *Atlas of Protein Side Chain Interactions,* IRL Press at Oxford University Press, Oxford
20. Miller, S., Janin, J., Lesk, A. M., and Chothia, C. (1987) *J. Mol. Biol.* **196,** 641–656
21. Lanyi, J. K. (1974) *Bacteriol. Rev.* **38,** 272–290
22. Baldacci, G., Guinet, F., Tillit, J., Zaccai, G., and de Recondo, A.-M. (1990) *Nucleic Acids Res.* **18,** 507–511
23. Dym, O., Mevarech, M., and Sussman, J. L. (1995) *Science* **267,** 1344–1346
24. Frolow, F., Harel, M., Sussman, J. L., Mevarech, M., and Shoham, M. (1996) *Nat. Struct. Biol.* **3,** 452–458
25. Bonete, M. J., Camacho, M. L., and Cadenas, E. (1987) *Int. J. Biochem.* **19,** 1149–1155
26. Cadenas, Q., and Engel, P. C. (1994) *Biochem. Mol. Biol. Int.* **33,** 785–792
27. Kraulis, P. J. (1991) *J. Appl. Crystallog.* **24,** 946–950
28. Nicholls, A., and Honig, B. (1991) *J. Comput. Chem.* **12,** 435–445
29. Ferrin, T. E., Huang, C. C., Jarvis, L. E., and Langridge, R. (1988) *J. Mol. Graphics* **6,** 13–27
30. Huang, C. C., Pettersen, E. F., Klein, T. E., Ferrin, T. E., and Langridge, R. (1991) *J. Mol. Graphics* **9,** 230–236

Biochemistry 1994, *33*, 8067–8073

Heme-Based Sensors, Exemplified by the Kinase FixL, Are a New Class of Heme Protein with Distinctive Ligand Binding and Autoxidation†

Marie A. Gilles-Gonzalez,* Gonzalo Gonzalez, and Max F. Perutz

Medical Research Council Laboratory of Molecular Biology, Cambridge CB2 2QH, United Kingdom

Laurent Kiger, Michael C. Marden, and Claude Poyart

INSERM Unité 299, 78, rue du Général Leclerc, F-94275, Le Kremlin-Bicêtre, France

Received March 8, 1994; Revised Manuscript Received May 2, 1994•

ABSTRACT: FixL's are chimeric heme protein kinases from symbiotic nitrogen-fixing *Rhizobia*. We have overexpressed three FixL variants in *Escherichia coli*. *Bradyrhizobium japonicum* FixL, a soluble dimeric protein, is the first full-length FixL to be purified. The other two proteins are soluble truncations of *Rhizobium meliloti* FixL, which is a membrane protein. One contains both heme and kinase domains and is dimeric; the other has only the heme domain and is monomeric. We find that all the FixL's bind oxygen and carbon monoxide non-cooperatively, with very low affinities due entirely to slow association rates. FixL P_{50}'s for oxygen are 17–76 mmHg. FixL's may sense nitric oxide and carbon monoxide in addition to oxygen, especially at the low oxygen pressures encountered *in vivo*. Autoxidation rates are about 50 times faster than that of sperm whale myoglobin. The carbon monoxide affinity of FixL's is about 300 times lower than that of myoglobin, resulting in the unusually low values of 7.5–17 for the partition constant, $M = P_{50}(O_2)/P_{50}(CO)$, between carbon monoxide and oxygen. Met-FixL's have their Soret absorption maximum at 395 nm instead of the typical 408 nm and a steep hydroxymet transition at pH ≥ 9.3; these properties indicate a pentacoordinated high-spin ferric heme and suggest a sterically hindered hydrophobic heme pocket lacking a distal (E7) histidine. FixL is the first member of a new class of heme proteins, the heme-based sensors, distinct from the oxygen carriers and electron transporters. We expect that some of the novel properties of FixL will be characteristic of the class.

Rhizobium meliloti FixL is a signal-transducing membrane protein that shuts down nitrogen fixation in response to oxygen (David et al., 1988). It is homologous to a diverse family of environmental sensors, ubiquitous in prokaryotes and recently also found in eukaryotes (Nixon et al., 1986; David et al., 1988; Chang et al., 1993). These proteins are modular; they consist of a nonconserved, usually N-terminal sensor domain that detects a specific environmental signal and a homologous transmitter domain with the kinase activity [reviewed by Stock et al. (1989) and Parkinson and Kofoid (1992)]. A genetically engineered soluble FixL, called FixL*, proved to be a kinase and heme protein with the characteristic oxy and deoxy spectra of Hb[1] (Gilles-Gonzalez et al., 1991). On the basis of homology, the transmitter domain is in the C-terminal two-thirds of FixL*; the N-terminal one-third is the sensor (Figure 1). FixL* kinase activity under nitrogen is about 7 times faster than in air (Gilles-Gonzalez et al., 1993). Deoxy-FixL*, the active form, transfers a phosphoryl group to the transcriptional activator FixJ, which enhances FixJ's activity and induces the expression of critical nitrogen fixation genes (Gilles-Gonzalez et al., 1991; Gilles-Gonzalez & Gonzalez, 1993; Agron et al., 1993).

We examined three FixL's that we overexpressed in *Escherichia coli* and purified (Figure 1). Two are new truncations of *R. meliloti* FixL; the other is the first FixL to be isolated from a different species and the first full-length FixL to be purified. All three are heme proteins. They are (1) RmFixLT, a more precise truncation of *R. meliloti* FixL that differs from FixL* only in the first 10 amino acids; (2) RmFixLH, a truncation containing only the sensor domain of *R. meliloti* FixL; and (3) BjFixL, a FixL from *Bradyrhizobium japonicum* that is soluble in its native form.

FixL's have no homology to known heme proteins (Vinogradov et al., 1993), and the structure of the heme binding pocket of FixL is still unknown. On the other hand, much is known about the relationship between structure, function, and spectral properties of a great variety of natural and genetically engineered Hb's (Perutz, 1989). We have tried to infer the nature of the heme environment of FixL's and to clarify the biological function of FixL's by comparing their ligand binding and absorption spectra to those of Hb's of known structure. FixL's bind oxygen non-cooperatively, and they differ dramatically from typical globins in their association rates with oxygen and carbon monoxide, susceptibility to autoxidation, and affinity of their ferric form for water.

† M.F.P. was supported by National Institutes of Health Grant HL31461 and the Medical Research Council. M.G.-G. was supported by the 1991 Forum Engelberg Prize to M.F.P. C.P., L.K., and M.M. were supported by funds from INSERM, la Direction des Recherches, Etudes et Techniques, the Faculté de Médecine Paris-Sud, and the Air Liquide Co.

* Corresponding Author. Tel. (0223)402289. Fax: (0223)213556.

• Abstract published in *Advance ACS Abstracts*, June 1, 1994.

[1] Abbreviations: BjFixL, *Bradyrhizobium japonicum* FixL; RmFixLT, *Rhizobium meliloti* FixL truncation with heme and kinase domains; RmFixLH, *Rhizobium meliloti* FixL truncation containing only the heme domain; Mb, myoglobin; Hb, hemoglobin; HbA, human adult hemoglobin; R, relaxed; $P(O_2)$, partial pressure of oxygen; P_{50}, partial pressure of a gaseous ligand at which the hemes are 50% saturated. The alphanumeric code (e.g., E7) refers to the positions of amino acids in helices and turns of myoglobin and hemoglobin (Perutz, 1970). The notation for myoglobin mutants is as follows: a one-letter code for the original amino acid followed by the alphanumeric code for the mutated position in parentheses, followed by a one-letter code for the replacing amino acid; e.g., H(E7)→V refers to the E7 position mutated from histidine to valine.

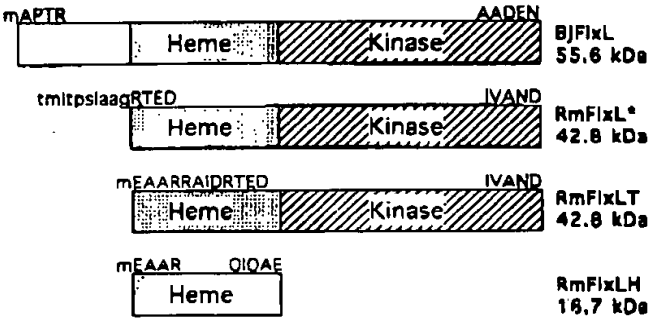


FIGURE 1: Schematic representations of FixL's. *B. japonicum* FixL (BjFixL), *R. meliloti* FixL heme domain (RmFixLH), *R. meliloti* truncated FixL containing heme and kinase domains (RmFixLT), and *R. meliloti* FixL* (RmFixL*), a previously studied truncation (Gilles-Gonzalez et al., 1991; Gilles-Gonzalez & Gonzalez, 1993), are shown. Monomer molecular masses, calculated from the amino acid sequence, are given. Terminal amino acid sequences are at the top of each diagram (David et al., 1988; Anthamatten & Hennecke, 1991); residues shown in lowercase type were introduced during cloning. The hatched kinase regions are highly homologous in sensors belonging to the two-component regulatory systems (Nixon et al., 1986). FixL's also share extensive homology in the stippled heme regions.

MATERIALS AND METHODS

Cloning. DNA manipulations were done according to manufacturers' recommendations and the methods compiled in Sambrook et al. (1989). The *B. japonicum fixL* gene was kindly provided by H. Hennecke on plasmid pRJ7349 (Anthamatten & Hennecke, 1991). The *B. japonicum fixL* gene was prepared from this plasmid in a polymerase chain reaction as a 1.5-kb fragment with an *NdeI* site at its 5′ end, overlapping the start codon, and a *Hin*dIII site at its 3′ end. Polymerase chain reactions were also used to prepare a 1.16-kb DNA fragment encoding RmFixLT and a 0.44-kb fragment encoding RmFixLH from plasmid pMW2 (Weinstein et al., 1992). Primers introduced an ATG start codon that overlaps with an *NdeI* site and added a *Hin*dIII site at the 3′ end. For RmFixLH, stop codons were added to the end of the coding sequence in all three frames.

Each of the three *NdeI–Hin*dIII fragments described above was cloned into expression vector pNKT (N. Komiyama and J. Avis, unpublished data). This pUC8-derived vector carries ampicillin resistance (Messing, 1983); the polylinker region is preceded by a *tac* promoter and two tandem ribosome binding sites. The plasmids containing the genes encoding BjFixL, RmFixLT, and RmFixLH are named pBL31, pRT51, and pRH61, respectively. Clones without polymerization error were assembled after verification of the *fixL* sequences by direct double-stranded plasmid sequencing (Sanger et al., 1977; Smith et al., 1979). Each protein was produced from *E. coli* strain TG1 (Gibson, 1984) transformed with the appropriate vector.

Expression and Purification. BjFixL and RmFixLT were expressed from TG1(pBL31) and TG1(pRT51), respectively, after induction with IPTG. The proteins were purified as described for *R. meliloti* FixL* (Gilles-Gonzalez et al., 1991).

RmFixLH was purified as follows. Induction with IPTG and preparation of cleared lysate of TG1(pRH61) were as described for *R. meliloti* FixL* (Gilles-Gonzalez et al., 1991). RmFixLH from a 25–50% saturated ammonium sulfate cut of the cleared lysate was further purified on a Sephacryl S-100 gel filtration column equilibrated in 5 mM Tris-HCl, pH 8.0, and 10 mM β-mercaptoethanol.

All column matrices were from Pharmacia. Throughout the purifications, FixL was kept at <4 °C and was assayed on the basis of 410-nm absorbance. Protein was assayed by the method of Bradford (Bio-Rad dye reagent concentrate), using bovine serum albumin as the standard. Gel filtration columns were calibrated with mixtures of molecular weight markers prior to the runs. A mixture of aldolase (158 kDa), albumin (67 kDa), and ovalbumin (43 kDa) was used for Sephacryl S-200 calibration. The Sephacryl S-100 column was calibrated with a mixture of albumin (67 kDa), ovalbumin (43 kDa), α-chymotrypsinogen (25 kDa), and ribonuclease (13.7 kDa).

Spectrophotometric Studies. Spectra (300–700 nm) of samples at 25 °C, in 20 mM phosphate buffer at pH 7, were recorded with an SLM Aminco DW2000 apparatus. Due to the fast autoxidation of FixL's, their met spectra could be obtained by keeping them for a few hours in air at 25 °C. The samples, in stoppered quartz cuvettes, were then deoxygenated with a stream of N_2; deoxy spectra were recorded after reduction with a 2-fold molar excess of sodium dithionite. To obtain oxy-FixL while avoiding autoxidation, one atmosphere of O_2 was layered over the deoxy sample at 5 °C; the solutions were equilibrated by shaking just before the spectra were recorded. CO spectra were obtained from samples equilibrated with 1 atm of CO. Heme content was measured from the protoheme pyridine hemochromogen (Appleby, 1978).

Autoxidation. The rate of autoxidation was measured at 37 °C on 10–20 μM protein in 75 mM potassium phosphate, pH 7.0, and 0.2 mM EDTA. Ferric FixL was reduced with crystalline sodium dithionite. Oxy-FixL, free of dithionite, was recovered from a small Sephadex G-25 (Pharmacia) column equilibrated with air-saturated 75 mM sodium phosphate, pH 7.0, and 0.2 mM EDTA. Alternatively, an anaerobic solution of met-FixL was reduced with a 2-fold molar excess of dithionite, and oxy-FixL was produced by equilibration of the ice-cooled sample with 1 atm of air or O_2. The oxy-FixL was warmed rapidly to 37 °C, and rates were measured by monitoring absorbance from 350 to 700 nm at intervals of 5–10 min. Both methods of preparation of oxy-FixL gave similar results. All time courses showed clear isosbestic points. Rates were calculated from the absorbance change at 577 nm.

pH Titrations. For each FixL, seven identical aliquots were adjusted to a desired pH and the same final concentration. Titrations were done on 10–20 μM protein in 0.1 M buffer containing 0.2 mM EDTA. Buffers were sodium phosphate for pH 6.5–8.0, glycine NaOH or sodium carbonate/bicarbonate for pH 9.2–10.7, and CAPS at pH 11.1. Visible spectra from 350 to 700 nm were recorded. Absorbance at 605 nm was a measure of transition to the hydroxyl met form.

O_2 Binding at Equilibrium. Binding curves were measured by a continuous method, with a Hemox-Analyzer (TCS Huntington Valley, Southampton, PA), as detailed elsewhere (Kister et al., 1986). Samples were at 25 °C in 50 mM Tris-HCl, pH 7.5, and 0.1 M NaCl containing 20 μg of catalase and 50 μM EDTA. Met-FixL was reduced with sodium dithionite and stripped of the reductant through a small G-25 column equilibrated with aerated buffer in the cold. The P_{50} and the index of cooperativity, n, were calculated from linear regression analysis of points between 80 and 20% O_2 saturation. For comparison, we measured the O_2 affinity of freshly prepared isolated α chains of human Hb in similar conditions but without the addition of dithionite.

CO Binding at Equilibrium. Deoxy-BjFixL was prepared by equilibrating met-BjFixL (3 mL of 20 μM in 13.3 mM phosphate buffer, pH 7) with 1 atm of N_2 at 37 °C in a stoppered tonometer and reducing the protein with sodium

O_2 and

dithioni
was ther
a strean
mmHg
saturate
saturate
of CO at
spectra.

Ligan
FixL pr
photolys
YAG la
phospha
reduced
or CO be
nm was r
was usec
unusuall

Ligan
for BjFi
stopped-
sample,
mM pote
of 10 μM
at 425 n
0.2 mM
compare

Accur
rate [k_o
replacem
plete oxy
rapid aut
for RmF
25 °C.

The k_o
to a meth
by Kiger
measure
allosteric
photodis
a mixture
O_2 may
CO. If t
can be es
is exposed
was at 4

RESULT

Expres
from DE
is similar
Gonzalez
estimates
respectiv
we found
molecula
per liter
mg for R
focusing
On the ba
FixL's c
monomer

dithionite. The deoxy spectrum was recorded. The sample was then equilibrated for 15 min at room temperature under a stream of 0.010 atm of CO, resulting in a P(CO) of 7.1 mmHg in the tonometer. The spectrum of partially CO-saturated FixL was recorded. The spectrum of 100% CO-saturated FixL was obtained after equilibration with 1 atm of CO at 25 °C. The P_{50}(CO) was estimated from these three spectra.

Ligand Association. The rates of association (k_{on}) of the FixL proteins with CO and O_2 were obtained after flash photolysis with a 10-ns 160-mJ pulse at 532 nm (Quantel YAG laser, France). The heme proteins (10 μM, in 20 mM phosphate buffer, pH 7, at 25 °C in 4-mm cuvettes) were reduced with sodium dithionite and exposed to 1 atm of O_2 or CO before the flash photodissociation. Absorbance at 436 nm was monitored; a low-intensity monochromatic light source was used to avoid photolysis of these samples, which have unusually slow ligand on-rates.

Ligand Dissociation. The CO dissociation rates [k_{off}(CO)] for BjFixL and RmFixLH at 25 °C were measured with a stopped-flow apparatus (Biologic, France) by mixing the sample, equilibrated with 0.1 atm (100 μM) of CO, with 2 mM potassium ferricyanide, resulting in final concentrations of 10 μM free CO and 1.8 mM ferricyanide. Detection was at 425 nm. Control experiments with the deoxy protein and 0.2 mM ferricyanide showed that oxidation is fast (<1 s) compared to CO dissociation kinetics.

Accurate and reproducible measurements of O_2 dissociation rate [$k_{off}(O_2)$] at 25 °C were not possible by O_2/CO replacement with the stopped-flow apparatus, due to incomplete oxygenation of the FixL's, even at 1 atm O_2, and their rapid autoxidation. The $k_{off}(O_2)$ value given in parentheses for RmFixLT was calculated from P_{50} and $k_{on}(O_2)$ values at 25 °C.

The $k_{off}(O_2)$ for BjFixL at 25 °C was measured according to a method described by Astatke et al. (1992) and modified by Kiger et al. (1993). This method allows O_2 binding measurements for low oxygen affinity or easily oxidized non-allosteric hemoproteins. The principle of this technique is to photodissociate a protein sample previously equilibrated with a mixture of O_2 and CO. After photodissociation of the CO, O_2 may rebind to the exposed heme; next O_2 is replaced by CO. If the gas mixture contains nonsaturating O_2, the P_{50} can be estimated. Oxidation is not a problem, as the heme is exposed to O_2 for less than 1 s. For all these studies, detection was at 436 nm.

RESULTS

Expression and Purification. RmFixLT and BjFixL eluted from DEAE-Sephacel around 200 mM NaCl, at pH 7.8. This is similar to the elution behavior of *R. meliloti* FixL* (Gilles-Gonzalez et al., 1991). The gel filtration molecular mass estimates of RmFixLT and BjFixL are 91 and 130 kDa, respectively, indicating that they are dimeric. By contrast, we found that RmFixLH is monomeric, with an estimated molecular mass of 21 kDa. Recoveries of >95% pure protein per liter of bacterial culture are 10–12 mg for BjFixL, 6–8 mg for RmFixLT, and 3–4 mg for RmFixLH. Isoelectric focusing verified the purity of our preparations (Figure 2). On the basis of protein and pyridine hemochromogen assays, FixL's contain one iron protoporphyrin IX per protein monomer (Falk, 1964; Appleby, 1978).

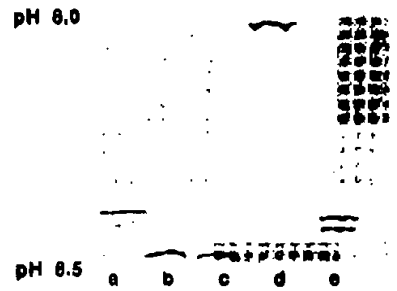


FIGURE 2: Isoelectric focusing of HbA (a, e), RmFixLH (b), RmFixLT (c), and BjFixL (d). The samples were focused in a gradient of pH 6–9 using an LKB system.

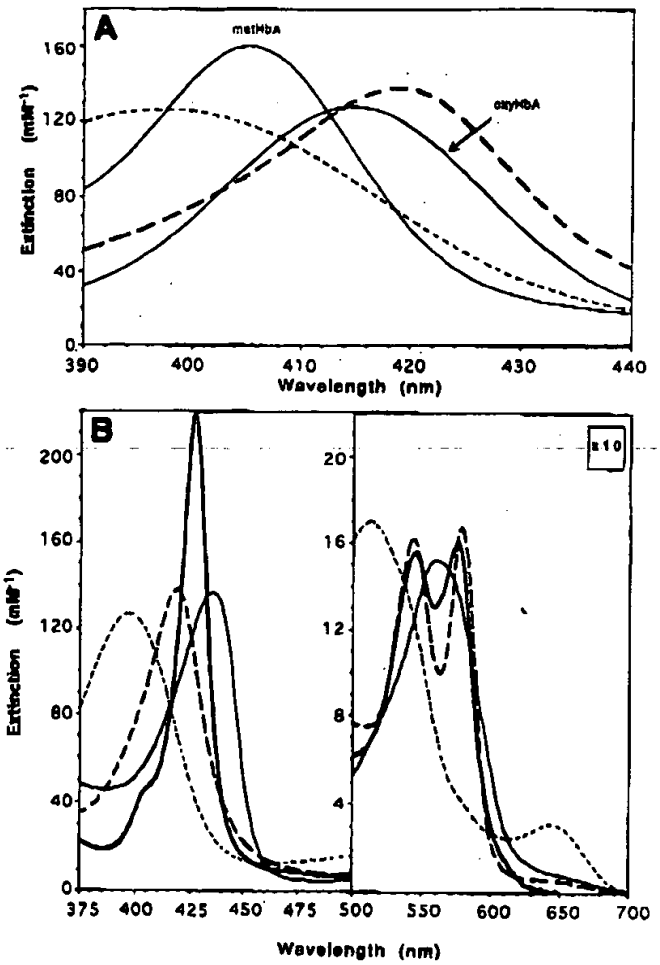


FIGURE 3: Absorption spectra of FixL proteins at pH 7.0. Panel A shows a comparison of the Soret bands of metHbA and oxyHbA (solid lines) to those of met-FixL (thin broken line) and oxy-FixL (thick broken line). Panel B shows the met (thin broken line), oxy (thick broken line), CO (thick solid line), and deoxy (thin solid line) spectra of RmFixLT. The absorption spectra of the *R. meliloti* and *B. japonicum* FixL's are nearly identical, except near 280 nm. Details of sample preparation are given in Materials and Methods.

Spectral Properties. Figure 3 shows the absorption spectra of the FixL's, which are nearly identical at 350–700 nm. Pure ferric BjFixL, RmFixLT, and RmFixLH have absorbance ratios of A_{280nm}/A_{395nm} = 0.69, 0.34, and 0.24, respectively (not shown). The oxy, deoxy, and carbonmonoxy spectra of FixL are similar to those of Hb, but the met spectra are markedly different (Figure 3) (Van Assendelft, 1970; Perutz et al., 1974). Hb, which exists below pH 8 predominantly as the aquomet form with water at the sixth coordination position, has an intense Soret band at 405 nm with an extinction

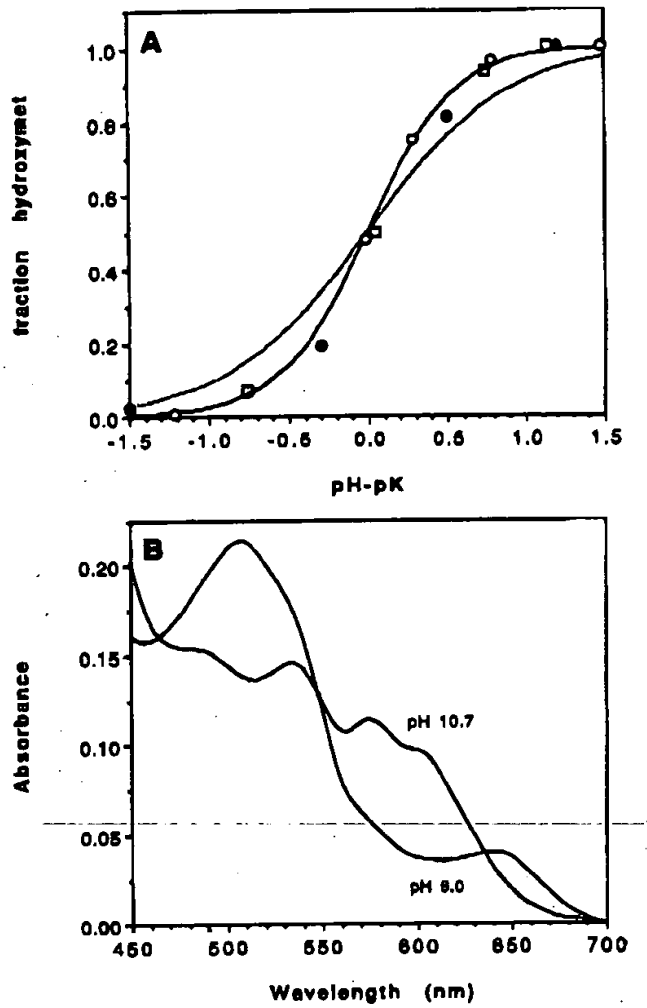


FIGURE 4: pH titration of met-FixL's. Panel A shows the pH titrations of the met forms of BjFixL (□), RmFixLT (●), and RmFixLH (○). The dotted line is the titration curve of sperm whale Mb; the solid line represents an average titration curve of the FixL's. Absorbance at 605 nm was a measure of the hydroxymet form. Panel B shows the absorption spectra of RmFixLT at pH 8.0 and 10.7. Details are in Materials and Methods.

coefficent of 161 mM^{-1} cm^{-1} heme, much higher than that of oxyHbA (131 mM^{-1} cm^{-1}) (Figure 3A). In contrast, the Soret band of met-FixL is blue-shifted to 395 nm with an extinction of 126 mM^{-1} cm^{-1}, lower than that of oxy-FixL; this indicates that met-FixL is pentacoordinated below pH 8 (Shikama & Matsuoka, 1989; Quillin et al., 1993). With increasing pH, the high-spin peak of Hb at 630 nm gradually disappears as the heme-linked water is titrated to a hydroxyl group. FixL's also undergo a transition from the high-spin "acid" spectrum to the low-spin alkaline spectrum, with pK_a's of 9.3 for RmFixLH and RmFixLT and 10 for BjFixL, but this transition is from a five-coordinated to a six-coordinated ferric iron (Figure 4). Hydroxide binds to the heme cooperatively, with Hill parameter $n = 1.5$.

Autoxidation. Clear isosbestic points occur at 409, 461, 535, and 600 nm for the decay of the oxy-FixL's to their corresponding met-FixL's. The half-lives of the oxy derivatives of BjFixL, RmFixLT, and RmFixLH are 15, 20, and 18 min, respectively, in air at pH 7, 37 °C (Table 1). This represents an autoxidation rate about 50 times faster than that of native sperm whale Mb or its mutant which has the distal histidine replaced by glutamine [H(E7)→Q] but 3–30 times slower than those of other Mb mutants lacking a distal histidine (Quillin et al., 1993). Springer et al. (1989) have reported even faster autoxidation rates, about 50 times those of the FixL's, for the sperm whale Mb mutants H(E7)→Y and H(E7)→D.

Table 1: Autoxidation of FixL Proteins in Air at pH 7, 37 °C

protein	k_{ox} (h^{-1})	half-life (min)	met Soret λ_{max} (nm)	oxy Soret λ_{max} (nm)
BjFixL	2.7	15	395	419
RmFixLT	1.9	22	395	418
RmFixLT[a]	2.1	20	395	418
RmFixLH	2.3	18	395	418
Aplysia Mb[b]			395	418
SW Mb[c]	0.06	720	409	418
SW Mb H(E7)→L[c]	10	4	393	

[a] One atmosphere of O_2 instead of air. [b] Shikama and Matsuoka (1989). [c] Sperm whale myoglobin (Quillin et al., 1993).

Oxygen and Carbon Monoxide Binding. FixL's bind oxygen non-cooperatively (Hill parameter $n = 1.0$) and have lower oxygen affinities than most natural Mb's or relaxed Hb (Table 2). Their P_{50}'s are 40–150 times the 0.3–0.5 mmHg typical of Mb, HbA (R), or its free α and β subunits. This unusually low oxygen affinity is due to slow oxygen association rates; the dissociation rates are comparable to those of globins (Table 2). BjFixL has a P_{50} of 76 mmHg, a value much higher than for the *R. meliloti* FixL's. The presence of a kinase domain in RmFixLT increases P_{50} to 27 mmHg from the 17 mmHg of RmFixLH. The rates of association of FixL with carbon monoxide are even more unusual: BjFixL has a rate 100 times slower than that of Mb and over 1000 times slower than that of HbA (R) (Table 2). The carbon monoxide dissociation rates are similar to those of Hb.

DISCUSSION

Typical metMb has water bound at the sixth coordination position of the ferric heme iron at low pH. Quillin et al. (1993) have compared the visible spectra and three-dimensional structures of a variety of Mb mutants. They found that hexacoordinated ferric heme iron has its Soret absorption maximum at 407–410 nm, while pentacoordinated metMb has it blue-shifted to 390–396 nm. Distal histidine, glutamine, or other hydrogen-bonding residues stabilize the heme-linked water. Hydrophobic distal residues, like phenylalanine, valine, or leucine, prevent binding of water to the sixth coordination site and result in a pentacoordinated ferric iron. A distal threonine results in partial occupancy of water and an intermediate spectral behavior. Below pH 9, the Soret absorptions of the met-FixL's are characteristic of a high-spin pentacoordinated heme iron (Figure 3A; Table 1). While in Mb or HbA the transition at high pH is from a water at the sixth coordination position to a hydroxyl, the FixL transition is a more abrupt one from a pentacoordinated to a hexacoordinated heme iron (Figure 4). The cooperativity of hydroxide binding, even in the monomeric RmFixLH, suggests that the affinity of the heme toward hydroxide is linked to deprotonation or hydroxide binding at another site in the monomer. The transition pH's of 9.3–10 for FixL's and ~8.5 for human HbA correspond to a free energy of hydroxymet formation that is 1.3–2.2 kcal mol^{-1} more positive in FixL's.

Despite the difficulties we encountered due to the high autoxidation rates of the FixL's, the oxygen affinities derived from our equilibrium measurements agree with those calcu-

Table 2: Ligand Binding Parameters of FixL Proteins at 25 °C[a]

ligand	protein	P_{50} (mmHg)	K ($\times 10^{-4}$ M)	k_{on} ($\times 10^{-4}$ M^{-1} s^{-1})	k_{off} (s^{-1})	M[b]
O_2	BjFixL	76	0.73	14.5	20	17
	RmFixLT	27	≤2.0	21.7	(11)[c]	
	RmFixLH	17	3.2	21.7	6.8	7.5
	Aplysia Mb[d]	2.65	21	1500	70	100
	SW Mb[e]	0.48	≤20	1400	12	23
	SW Mb H(E7)→L[e]	23	2.4	9800	4100	48 000
	human HbA (R)	0.3	80	4000	50	300
CO	BjFixL	4.3	10	0.5	0.045	
	RmFixLT			1.2		
	RmFixLH	2.2	20	1.7	0.083	
	Aplysia Mb[d]	0.02	3000	50	0.02	
	SW Mb[e]	0.028	2700	51	0.019	
	SW Mb H(E7)→L[e]	0.0007	110 000	2600	0.024	
	human HbA (R)	0.001	50 000	600	0.013	

[a] Details in Materials and Methods. [b] Values calculated from $P_{50}(O_2)/P_{50}(CO)$. [c] This value is calculated from K and k_{on}. [d] Wittenberg et al. (1965, 1972). [e] Sperm whale myoglobin (Springer et al., 1989; Quillin et al., 1993).

Table 3: FixL Oxygenated at Various Oxygen Pressures

environment	$P(O_2)$ (mmHg)	% oxy-BjFixL	% oxy-RmFixLT	% oxy-RmFixLH
air	152	67	85	90
microaerobic culture	20	21	43	54
nodule	0.010	0.013	0.037	0.059

lated from our rate constants. They show that the oxygen affinities of FixL's are lower than those of almost all natural oxygen carriers. Even in air, RmFixLT contains 15% of the deoxy form, which is sufficient to account for the residual kinase activity we had observed (Gilles-Gonzalez et al., 1993). FixL-dependent induction of nitrogen fixation genes in free-living wild-type *Rhizobia* occurs even at partial pressures of oxygen as high as 20 mmHg (Ditta et al., 1987). Such induction is consistent with the high P_{50} of RmFixLT (Table 3) and suggests that this truncated FixL, *in vitro*, is a good model for membrane-bound wild-type FixL in *R. meliloti*. RmFixLT would be 99.96% deoxygenated and therefore have maximal kinase activity in nitrogen-fixing nodules, where the partial pressure of oxygen is about 10 μmHg (Table 3) (Appleby, 1969; Wittenberg et al., 1972). The oxygen affinity of RmFixLH is 60% greater than that of RmFixLT, perhaps because RmFixLH consists solely of a monomeric heme domain, while RmFixLT is a dimer containing heme and kinase domains in each subunit and is therefore probably more rigid (Table 2; Figure 1).

The kinase and heme domains of *B. japonicum* and *R. meliloti* FixL's are highly homologous, but not their N-terminal domains (Figure 1) (Anthamatten & Hennecke, 1991). The N-terminal domain of *R. meliloti* FixL contains several transmembrane segments, while the BjFixL N-terminal domain has none (Lois et al., 1993). BjFixL is essential for nitrogen fixation in *B. japonicum*, but its regulatory target is unknown; unlike *R. meliloti* FixL, it does not regulate nitrogenase production (Anthamatten & Hennecke, 1991). The coordinated expression of more than 20 gene products is required for nitrogen fixation. FixL's would begin to function as kinases at relatively high oxygen pressure, possibly inducing genes involved in the transition to and from the nearly anaerobic nodule environment (Table 2). The *fixNOPQ* gene cluster, which encodes the alternative oxidase complex of *B. japonicum* required to support bacterial respiration in nodules, is an excellent candidate (Preisig et al., 1993). FixL proteins may sense other heme ligands in addition to oxygen. In the nearly anaerobic environment of nodules, oxygen binding would not interfere with the detection of other ferrous heme ligands such as nitric oxide or carbon monoxide, even at very low concentrations.

The low oxygen affinity of FixL's is due entirely to a very slow association rate (Table 2). On the other hand, Mb mutants that lack heme-linked water in their met forms, due to replacement of the distal histidine by either leucine, valine, or phenylalanine, have off-rates for oxygen accelerated 300-, 460- and 700-fold, respectively. Surprisingly, the on-rates are also accelerated, although this is only by 6, 7, and 5 times, respectively. The replacements all increased the affinity for carbon monoxide, due to rises in on-rates outweighing smaller rises in off-rates. Their on-rates are raised probably because a water molecule, hydrogen-bonded to the distal histidine (but not bound to the ferrous iron) in native deoxyMb, presents a kinetic barrier to oxygen binding; this water molecule is absent when the distal histidine has been replaced by a hydrophobic residue. All these replacements in Mb raised the value of M between 400- and 1000-fold (Quillin et al., 1993), whereas FixL has a very low M. In *Aplysia limacina* Mb, the distal histidine is replaced by a valine, so that its ferric form also lacks a heme-linked water molecule; but unlike FixL, it has a rather high oxygen affinity due to a high on-rate and a slow off-rate (Table 2) (Conti et al., 1993; Wittenberg et al., 1965, 1972). The former is due to ready access to its open heme pocket, and the latter, to the ability of its arginine E10 to swing into the heme pocket and donate hydrogen bonds to electronegative heme ligands.

There is no model that explains the low affinities for both oxygen and carbon monoxide, caused almost entirely by slow on-rates. In their extensive experiments on the effects of proximal and distal hindrance in the synthetic picket fence heme complex, Collman, Gibson, and their colleagues (1983) found that steric hindrance on either side of the heme increased the off-rates for oxygen but left the on-rates almost unchanged. Proximal steric hindrance slowed the on-rates for carbon monoxide and accelerated the off-rates about equally. Distal hindrance slowed the on-rate for carbon monoxide and left the off-rate unchanged. This did lower M, but even the lowest value is still 100 times higher than that of FixL (Collman et al., 1983).

Autoxidation of Hb is often attributed to dissociation of superoxide ion (O_2^-) from oxyHb. Interesting new light on the mechanism that prevails in the absence of a distal hydrogen bond donor to the bound oxygen has emerged from a study of the autoxidation of the iron cyclidenes pictured in Figure 5A. Dickerson et al. synthesized a series of such bridged cyclidenes in which R1 is an aliphatic chain varying in length

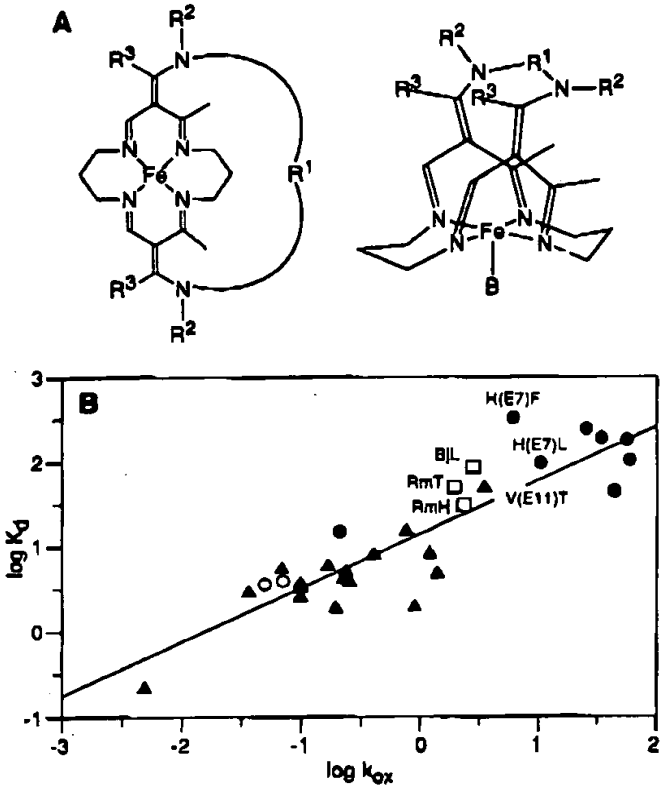


FIGURE 5: Relationship of FixL autoxidation to synthetic iron(II) cyclidene models and myoglobin variants. Panel A shows the structure of the bridged cyclidene [$R^1 = (CH_2)_4$] that most closely resembles FixL in its autoxidation (Dickerson et al., 1993). Panel B is a comparison of the autoxidation rates (h^{-1}) and oxygen equilibrium dissociation constants (μM) for FixL's and for myoglobins from sperm whale and pig: FixL's (□), wild-type Mb (○), histidine (E7) Mb mutants (●), other mutations of highly conserved Mb residues (▲). This plot is based on Figure 7 of Brantley et al. (1993) and Mb data from Tables II–V of the same article. BjFixL, RmFixLT, and RmFixLH are shortened to BjL, RmT, and RmH, respectively.

from $(CH_2)_3$ to $(CH_2)_5$ and above (Dickerson et al., 1993). $(CH_2)_3$ does not bind either O_2 or CO; $(CH_2)_4$ combines with the two gases with low affinity. Thereafter, the affinity rises with increasing length of the aliphatic chain. The oxidation rate of $(CH_2)_3$ is fast and follows simple first-order dependence on $P(O_2)$, while according to the superoxide dissociation mechanism it should not autoxidize at all. The authors argue that such first-order dependence is consistent only with an "outer shell mechanism", i.e., electrons being transferred from iron to oxygen colliding with the outer shell of the complex. The autoxidation rate of $(CH_2)_4$ reaches saturation at $P(O_2)$ > 60 mmHg, when oxygen combining with and protecting the ferrous iron competes successfully with electron transfer to oxygen. $(CH_2)_5$ shows a steep rise in autoxidation rate up to $P(O_2)$ = 10 mmHg and then a steep fall, with the rate approaching zero near $P(O_2)$ = 100 mmHg where protection by the bound oxygen dominates.

The mechanism of Dickerson et al. predicts that autoxidation rate will be inversely related to oxygen affinity, since the deoxy form is the reactive species. It further predicts that heme pockets in which absence of hydrogen bonding disfavors occupancy by water will show the autoxidation dependence on $P(O_2)$ of the type described above for $(CH_2)_4$-bridged cyclidenes. The approximately linear relationship between the log of the autoxidation rate and the log of the oxygen dissociation constant of Mb mutants, observed by Brantley et al. (1993), supports the Dickerson model. FixL fits very well on this plot (Figure 5B). When the distal histidine in myoglobin is replaced by a valine, its autoxidation rate varies with the partial pressure of oxygen in much the same way as that of the $(CH_2)_4$-bridged cyclidene (Brantley et al., 1993; Quillin et al., 1993). RmFixLT, which resembles the H(E7)→V mutant, in the absence of a water molecule linked to its ferric iron, has the same elevated autoxidation rate in air as at 1 atm of oxygen, again fitting this model (Table 1). The rates of autoxidation of FixL's and the bridged cyclidene that does not bind oxygen are comparable: 1.9–2.7 versus 0.9 h^{-1}.

There is little evolutionary pressure to evolve resistance to oxidation in the extremely reducing, nearly anaerobic environment of a nitrogen-fixing nodule. Despite the high autoxidation rates of FixL's (Table 1), even mildly reducing conditions protect their heme irons from oxidation. In 10 mM β-mercaptoethanol, in air, ferric *R. meliloti* FixL is reduced to oxy-FixL and does not autoxidize even after two days at room temperature. There is therefore no need for reductases of the kind present in red cells and muscle tissue.

What can we infer about the heme environment? The absorption spectra of deoxy-, oxy-, carbonmonoxy-, and met-FixL leave no doubt that the proximal heme ligand is a histidine, since heme proteins with cysteine, methionine, or tyrosine as proximal ligands show very different absorption spectra. The very low oxygen affinity and the absence of a water molecule at the sixth coordination position of the ferric heme are indicative of nonpolar residues on the distal side of the heme. The low association rates for both oxygen and carbon monoxide suggest that these residues hinder ligand binding. There is one highly conserved histidine in RmFixLH that is likely to be the proximal ligand. In the myoglobin sequence, the axial ligands are separated by 20–30 residues. Either of two phenylalanines that are about this distance from the conserved histidine might be the distal residue.

FixL proteins are the only known members of a third class of heme proteins, which we are calling the heme-based sensors; they are distinct from oxygen carriers or electron transporters. As such, the uniqueness of their properties is not surprising. We predict that other heme-based sensors exist, which sense not only oxygen but other heme ligands such as nitric oxide, and that these sensors will share some of the novel characteristics of FixL. Likely candidates include the endothelial nitric oxide sensors that regulate blood pressure and the sensors that regulate hypoxia-induced genes such as the erythropoeitin gene and the Fos and Jun protooncogenes (Blanchard et al., 1992; Webster et al., 1993; Wang & Semenza, 1993).

ACKNOWLEDGMENT

We thank Denise Anthamatten, Hauke Hennecke, Noburu Komiyama, and Johanna Avis for their gifts of plasmids. We gratefully acknowledge the expert technical assistance of Brigitte Bohn, Genevieve Caron, N. Griffon, and Jean Kister.

REFERENCES

Agron, P. G., Ditta, G. S., & Helinski, D. R. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 3506–3510.

Anthamatten, D., & Hennecke, H. (1991) *MGG, Mol. Gen. Genet. 225*, 38–48.

Appleby, C. A. (1969) *Biochim. Biophys. Acta 188*, 222–229.

Appleby, C. A. (1978) *Methods Enzymol. 52*, 157–166.

Astatke, M., McGee, W. A., & Parkhurst, L. (1992) *Comp. Biochem. Physiol. 101B*, 683–688.

Blanchard, K. L., Acquaviva, A. M., Galson, D. L., & Bunn, H. F. (1992) *Mol. Cell. Biol. 12*, 5373–5385.

Brantley, R. E., Smerdon, S. J., Wilkinson, A. J., Singleton, E. W., & Olson, J. S. (1993) *J. Biol. Chem. 268*, 6995–7010.
Chang, C., Kwok, S. F., Bleecker, A. B., & Meyerowitz, E. M. (1993) *Science 262*, 539–544.
Collman, J. P., Brauman, J. I., Iverson, B. L., Sessler, J. L., Morris, R. M., & Gibson, Q. H. (1983) *J. Am. Chem. Soc. 105*, 3052–3064.
Conti, E., Moser, C., Rizzi, M., Mattevi, A., Lionetti, C., Coda, A., Ascenzi, P., Brunori, M., & Bolognesi, M. (1993) *J. Mol. Biol. 233*, 498–508.
David, M., Daveran, M.-L., Batut, J., Dedieu, A., Domergue, O., Ghai, J., Hertig, C., Boistard, P., & Kahn, D. (1988) *Cell 54*, 671–683.
Dickerson, L. D., Sauer-Masarwa, A., Herron, N., Fendrick, C. M., & Busch, D. H. (1993) *J. Am. Chem. Soc. 115*, 3623–3626.
Ditta, G., Virts, E., Palomares, A., & Kim, C.-H. (1987) *J. Bacteriol. 169*, 3217–3223.
Falk, J. E. (1964) *Prophyrins and metalloporphyrins*, p 240, Elsevier Publishing Co., London.
Gibson, T. J. (1984) Ph.D. Thesis, Cambridge University, England.
Gilles-Gonzalez, M. A., & Gonzalez, G. (1993) *J. Biol. Chem. 268*, 16293–16297.
Gilles-Gonzalez, M. A., Ditta, G. D., & Helinski, D. R. (1991) *Nature 350*, 170–172.
Kiger, L., Poyart, C., & Marden, M. C. (1993) *Biophys. J. 65*, 1–9.
Kister, J., Poyart, C., & Edelstein, S. J. (1987) *J. Biol. Chem. 262*, 12085–12091.
Lois, A. F., Ditta, G. S., & Helinski, D. R. (1993) *J. Bacteriol. 175*, 1103–1109.
Messing, J. (1983) *Methods Enzymol. 101*, 20–78.
Nixon, B. T., Ronson, C. W., & Ausubel, F. M. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 7850–7854.
Parkinson, J. S., & Kofoid, E. C. (1992) *Annu. Rev. Genet. 26*, 71–112.
Perutz, M. F. (1970) *Nature 228*, 726–739.
Perutz, M. F., Heidner, E. J., Ladner, J. E., Bellestone, J. P., Ho, C., & Slade, E. F. (1974) *Biochemistry 13*, 2187–2200.
Perutz, M. F. (1989) In *Mechanisms of Cooperativity and Allosteric Regulation in Proteins*, pp 1–27, University Press, Cambridge.
Preisig, O., Anthamatten, D., & Hennecke, H. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 3309–3313.
Quillin, M. L., Arduini, R. M., Olson, J. S., & Phillips, G. N (1993) *J. Mol. Biol. 234*, 140–155.
Sambrook, J., Fritsch, E. F., & Maniatis, T. (1989) *Molecular cloning: A laboratory manual*, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Sanger, F., Nicklen, S., & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463–5467.
Shikama, K., & Matsuoka, A. (1989) *J. Mol. Biol. 209*, 489–491.
Smith, M., Leung, D. W., Gillam, S., Astell, C. R., Montgomery, D. L., & Hall, B. D. (1979) *Cell 16*, 753–761.
Springer, B. A., Egeberg, K. D., Sligar, S. G., Rohlfs, R. J., Mathews, A. J., & Olson, J. S. (1989) *J. Biol. Chem. 264*, 3057–3060.
Stock, J. B., Ninfa, A. J., & Stock, A. M. (1989) *Microbiol. Rev. 53*, 450–490.
Van Assendelft, O. W. (1970) *Spectrophotometry of haemoblogin derivatives*, Chapter 3, pp 47–73, Royal Vangorcum Ltd., Assen, the Netherlands.
Vinogradov, S. N., Walz, D. A., Pohajdak, B., Moens, L., Kapp, O. H., Suzuki, T., & Trotman, C. N. A. (1993) *Comp. Biochem. Physiol. 106B*, 1–26.
Wang, G. L., & Semenza, G. L. (1993) *J. Biol. Chem. 268*, 21513–21518.
Webster, K. A., Discher, D. J., & Bishopric, N. H. (1993) *J. Biol. Chem. 268*, 16852–16858.
Weinstein, M., Lois, A. F., Monson, E. K., Ditta, G. S., & Helinski, D. R. (1992) *Mol. Microbiol. 6*, 2041–2049.
Wittenberg, B. A., Brunori, M., Antonini, E., Wittenberg, J. B., & Wyman, J. (1965) *Arch. Biochem. Biophys. 111*, 576–579.
Wittenberg, J. B., Appleby, C. A., & Wittenberg, B. A. (1972) *J. Biol. Chem. 247*, 527–531.